



DISEÑO DEL LABORATORIO DE BASES DE DATOS PARA EL PROGRAMA DE
INGENIERÍA DE SISTEMAS DE LA UNIVERSIDAD CATÓLICA DE COLOMBIA

JOSÉ RODRIGO CHARRY LEMUS
EDWARD FELIPE CUELLAR LÓPEZ

UNIVERSIDAD CATÓLICA DE COLOMBIA
FACULTAD DE INGENIERÍA
PROGRAMA DE INGENIERÍA DE SISTEMAS
BOGOTÁ D.C.
2014



DISEÑO DEL LABORATORIO DE BASES DE DATOS PARA EL PROGRAMA DE
INGENIERÍA DE SISTEMAS DE LA UNIVERSIDAD CATÓLICA DE COLOMBIA

JOSÉ RODRIGO CHARRY LEMUS
EDWARD FELIPE CUELLAR LÓPEZ

Director
LUIS FELIPE HERRERA QUINTERO Ph. D.
Director de Programa

UNIVERSIDAD CATÓLICA DE COLOMBIA
FACULTAD DE INGENIERÍA
PROGRAMA DE INGENIERÍA DE SISTEMAS
BOGOTÁ D.C.
2014



Atribución-CompartirIgual 2.5 Colombia (CC BY-SA 2.5 CO)

Este es un resumen legible por humanos (y no un sustituto) de la [licencia](#).

[Advertencia](#)



Usted es libre para:



Compartir — copiar y redistribuir el material en cualquier medio o formato

Adaptar — remezclar, transformar y crear a partir del material

Para cualquier propósito, incluso comercialmente

El licenciante no puede revocar estas libertades en tanto usted siga los términos de la licencia

Bajo los siguientes términos:



Atribución — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.



CompartirIgual — Si usted mezcla, transforma o crea nuevo material a partir de esta obra, usted podrá distribuir su contribución siempre que utilice la misma licencia que la obra original.

No hay restricciones adicionales — Usted no puede aplicar términos legales ni medidas tecnológicas que restrinjan legalmente a otros hacer cualquier uso permitido por la licencia.

Nota de Aceptación

Aprobado por el comité de grado en cumplimiento de los requisitos exigidos por la Facultad de Ingeniería y la Universidad Católica de Colombia para optar al título de ingenieros de Ingeniero de Sistemas

Luis Felipe Herrera Quintero
Director

Carlos Alberto Pulido Leal
Revisor Metodológico

Bogotá D.C 18 de Noviembre de 2014

AGRADECIMIENTOS

Este trabajo de grado, forma parte de todo un proceso que se forjó durante años, para optar por el título de Ingeniero de Sistemas, todo este período requirió de constantes esfuerzos y sacrificios. En reconocimiento de alcanzar este logro queremos agradecer en primera medida a Dios por permitirnos crecer cada día como ser humano e ir alcanzando día tras día nuevas metas; a nuestros padres que estuvieron en todo este periodo apoyándonos incondicionalmente, a la Universidad Católica de Colombia por hacernos participe en este plan de mejoramiento, por contribuir con principios y valores en la formación personal y a todos los docentes que hicieron parte de este proceso de formación.

“A Dios, por ser mi guía y mi fortaleza, a mis padres, mi hermano por su apoyo incondicional, a mi mujer y amigos por su colaboración y hacer parte de este proceso y a los docentes por formarme y compartir sus conocimientos.” José Rodrigo Charry Lemus

“A Dios, por acompañarme y guiarme a lo largo de mi carrera, a mis padres Crisanto y Stella, por apoyarme en todo momento, por los valores inculcados, por la oportunidad de tener un excelente educación, a mi familia, amigos y a la Universidad por ser parte de este proceso tan importante en mi vida.” Edward Felipe Cuellar López

Finalmente queremos expresar nuestro más sincero agradecimiento a nuestro director de trabajo de grado, quien nos retribuyo su conocimiento en este proyecto y que además nos guio y superviso en la misma. A todas estas personas que aportaron y creyeron en nuestros capacidades, muchas gracias.

CONTENIDO

	Pág.
INTRODUCCIÓN	25
1. ANTECEDENTES	27
2. JUSTIFICACIÓN	30
3. PLANTEAMIENTO DEL PROBLEMA	32
4. OBJETIVOS	34
4.1 OBJETIVO GENERAL	34
4.2 OBJETIVOS ESPECÍFICOS	34
5. HIPÓTESIS	35
6. MARCO CONCEPTUAL	36
6.1 INFRAESTRUCTURA	36
6.2 HARDWARE	36
6.3 SOFTWARE	36
6.4 REDES	36
6.4.1 Clasificación de las redes	37
6.5 VIRTUALIZACIÓN	37
6.6 ACCESO REMOTO	37
6.7 LABORATORIO	38
6.8 LABORATORIOS VIRTUALES (LV)	38
6.9 LABORATORIO REMOTO (LR)	38
6.10 CABLEADO ESTRUCTURADO	39
6.11 RACK	39
6.12 BASES DE DATOS	39
6.13 BASE DE DATOS ESTRUCTURADO	40
6.14 BASE DE DATOS NO ESTRUCTURADOS	40
6.15 GESTOR DE BASES DE DATOS (SGBD)	40
6.16 DATAMINIG	40
6.17 BIG DATA	41
6.18 MAP REDUCE	41

6.19	DATA WAREHOUSE	41
6.20	HERRAMIENTAS REQUERIDAS EN EL LABORATORIO DE BASES DE DATOS	42
6.20.1	Sistema gestor de bases de datos	42
6.20.1	Modelado de datos	43
6.20.2	Big data	45
6.20.3	Data warehouse	50
6.20.4	Data mining	52
6.20.5	Virtualización.	54
6.20.6	Acceso remoto al laboratorio de bases de datos	60
7.	METODOLOGÍA	62
7.1	FASE DEL MÉTODO HIPOTÉTICO-DEDUCTIVO	62
7.1.1	Observación.	62
7.1.2	Formulación de la hipótesis.	63
7.1.3	Deducción de las consecuencias de la hipótesis.	63
7.1.4	Contrastación o verificación de la hipótesis.	63
8.	CRONOGRAMA	64
9.	ALCANCE Y LIMITACIONES	65
9.1	ALCANCE	65
9.2	LIMITACIONES	65
9.3	ESPACIO	65
9.4	TIEMPO	65
10.	ESTRATEGIAS DE COMUNICACIÓN	66
11.	PRESUPUESTO	67
12.	LEVANTAMIENTO DEL ESTADO ACTUAL DEL LABORATORIO DE BASES DE DATOS	68
12.1	HARDWARE Y SOFTWARE ACTUAL	72
12.1.1	Características de equipos de cómputo	72
12.2	CARACTERÍSTICAS DE SERVIDORES	72
12.3	SOFTWARE ACTUAL	76

12.4	USO DE LAS SALAS	77
12.4.1	Prácticas dirigidas	78
12.4.2	Prácticas libres	79
13.	LEVANTAMIENTO DE REQUERIMIENTOS DEL LABORATORIO DE BASE DE DATOS	83
13.1	ENCUESTA GENERAL REALIZADAS A LOS DOCENTES	84
13.1.1	¿Cuáles materias, según su enfoque académico, requieren de un incremento en la infraestructura de ti, para completar el desarrollo de las prácticas actuales?	84
13.1.2	¿Cree usted que los laboratorios requiere de accesos remotos para la realización de las prácticas?	85
13.1.3	¿En los laboratorios se cuenta con la posibilidad de instalar software que permitan el desarrollo de las prácticas en diferentes ambientes?	86
13.1.4	¿Los laboratorios cuentan con componentes de acceso remoto que amplien la disponibilidad, y apoyen el desarrollo de prácticas tanto dirigidas como libres?	87
13.1.5	¿Cuenta con los permisos de usuario necesarios para la realización de las prácticas en su totalidad?	88
13.1.6	¿Los laboratorios cuentan con el componente de virtualización?	89
13.1.7	Teniendo presente la disposición actual de las salas, bajo su criterio ¿cree que se necesita un mayor espacio para el número promedio de estudiantes que acceden a este para realizar las prácticas?	90
13.1.8	¿Los laboratorios cuentan con modalidades de acceso para estudiantes en condición de discapacidad?	91
13.1.9	¿Los laboratorios cuentan con la posibilidad de instalar, des instalar, y configurar software de acuerdo a necesidades de las asignaturas?	92
13.1.10	¿Considera que el hardware de la universidad satisface las necesidades de las asignaturas tratadas en esta encuesta?	93
13.2	ENCUESTA ESPECIFICA REALIZADA A LOS DOCENTES A FINES CON EL LABORATORIO DE BASES DE DATOS	94
13.2.1	¿Para cuáles de las unidades temáticas se presenta dificultad al realizar las prácticas propuestas en el currículo de la materia?	94

13.2.2 ¿Cree usted que el software actual de modelado de datos cumple con el propósito de las prácticas?	95
13.2.3 ¿Considera usted que el software actual de gestor de base de datos cumple con el propósito de las prácticas?	96
13.2.4 ¿Cree usted que el software actual, cumple con el propósito de las prácticas?	97
13.2.5 ¿Considera usted apropiado la implementación de tecnologías de virtualización en apoyo a las prácticas de laboratorio de bases de datos de forma remota?	98
13.2.6 ¿ Con los recursos actuales que se cuenta en los laboratorios es posible que usted realice prácticas en big-data?	99
13.2.7 ¿Con los recursos actuales que se cuenta en los laboratorios es posible que usted realice prácticas en data warehouse?	100
13.3 ESPECIFICACIÓN DE REQUERIMIENTOS	101
14. DISEÑO DEL LABORATORIO DE BASES DE DATOS	107
14.1 SERVICIOS	108
14.2 HARDWARE	111
14.2.1 Racks.	111
14.2.2 Servidores.	112
14.2.3 Acondicionamiento Térmico.	113
14.2.4 Equipos de cómputo.	115
14.2.5 Tablero interactivo.	116
14.3 SOFTWARE	116
14.4 INFRAESTRUCTURA	117
14.4.1 Adecuaciones físicas.	118
14.4.2 Área de Servidores.	120
14.4.3 Piso falso.	121
14.4.4 Acondicionamiento Térmico.	122
14.4.5 Sistema de seguridad y control.	122
14.5 ESPACIO DE TRABAJO	127
14.5.1 Espacio y movilidad.	127

15.	VALIDACIÓN DEL DISEÑO DEL LABORATORIO	135
15.1	VALIDACIÓN DE SERVICIOS	136
15.1.1	Servicio de acceso remoto.	137
15.1.2	Servicio de virtualización.	137
15.1.3	Servicio de permisos administrativos a usuarios.	139
15.1.4	Servicio de instalación de software.	140
15.1.5	Servicio de componentes de enfoque específico.	140
15.1.6	Servicio de práctica libre y dirigida.	140
16.	COSTEO DEL PROYECTO	143
17.	CONCLUSIONES Y APORTES	144
17.1	CONCLUSIONES	144
17.2	APORTES	145
18.	BIBLIOGRAFÍA	146
19.	ANEXOS	148

LISTA DE TABLAS

	Pág.
Tabla 1. Cronograma de actividades	64
Tabla 2. Talento humano	67
Tabla 3. Servicios	67
Tabla 4. Presupuesto global del proyecto	67
Tabla 5. Áreas del conocimiento	69
Tabla 6. Asignaturas área administración de la información	69
Tabla 7. Capacidad y disponibilidad salas	71
Tabla 8. Características equipos de cómputo	72
Tabla 9. Características servidor controlador de dominio	73
Tabla 10. Características servidor espejo	73
Tabla 11. Características servidor web apache	73
Tabla 12. Características servidor clúster	74
Tabla 13. Características servidor base de datos	74
Tabla 14. Análisis de uso de las sala 2 y la sala 4	82

Tabla 15. Requerimiento 1	101
Tabla 16. Requerimiento 2	102
Tabla 17. Requerimiento 3	102
Tabla 18. Requerimiento 4	103
Tabla 19. Requerimiento 5	103
Tabla 20. Requerimiento 6	104
Tabla 21. Requerimiento 7	104
Tabla 22. Requerimiento 8	105
Tabla 23. Requerimiento 9	105
Tabla 24. Requerimiento 10	106
Tabla 25. Rack del laboratorio de bases de datos	111
Tabla 26. Servidor rack del laboratorio de bases de datos	113
Tabla 27. Aire acondicionado de precisión del laboratorio de bases de datos	114
Tabla 28. Equipos de cómputo del laboratorio de bases de datos	115
Tabla 29. Tablero interactivo del laboratorio de bases de datos	116

Tabla 30. Ejemplos de distribución de luminarias en alumbrado general	118
Tabla 31. Soporte para el servicio de acceso remoto	137
Tabla 32. Soporte para el servicio de virtualización	138
Tabla 33. Soporte para el servicio de permisos administrativos a usuarios	139
Tabla 34. Soporte para el servicio de componentes de enfoque específico	140
Tabla 35. Matriz análisis de brecha	142
Tabla 36. Costeo del proyecto	143

LISTA DE FIGURAS

	Pág.
Figura 1. Ejemplo de un esquema json	46
Figura 2. Arquitectura de solr trabajando con hadoop.	47
Figura 3. Componentes de un agente de flume	49
Figura 4. Servidor con virtualización	55
Figura 5. Administración de imágenes con vmware	58
Figura 6. Esquema de acceso remoto	61
Figura 7. Método hipotético-deductivo	62
Figura 8. Mapa laboratorio de sistemas	71
Figura 9. Servidores U. Católica	75
Figura 10. Piso falso - sala informática	75
Figura 11. Sistema de detección de incendios	76
Figura 12. Servidores sala informática de la Universidad Católica de Colombia	76
Figura 13. Software disponible en salas	77
Figura 14. Número total usuarios prácticas dirigidas	78

Figura 15. Prácticas libres – Enero	79
Figura 16. Prácticas libres - Febrero	80
Figura 17. Prácticas libres – Marzo	80
Figura 18. Prácticas libres - Abril	81
Figura 19. Prácticas libres – Mayo	81
Figura 20. Pregunta #1 desarrollada en la encuesta general	84
Figura 21. Pregunta #2 desarrollada en la encuesta general	85
Figura 22. Pregunta #3 desarrollada en la encuesta general	86
Figura 23. Pregunta #4 desarrollada en la encuesta general	87
Figura 24. Pregunta #5 desarrollada en la encuesta general	88
Figura 25. Pregunta #6 desarrollada en la encuesta general	89
Figura 26. Pregunta #7 desarrollada en la encuesta general	90
Figura 27. Pregunta #8 desarrollada en la encuesta general	91
Figura 28. Pregunta #9 desarrollada en la encuesta general	92
Figura 29. Pregunta #10 desarrollada en la encuesta general	93

Figura 30. Pregunta #1 desarrollada en la encuesta específica	94
Figura 31. Pregunta #2 desarrollada en la encuesta específica	95
Figura 32. Pregunta #3 desarrollada en la encuesta específica	96
Figura 33. Pregunta #4 desarrollada en la encuesta específica	97
Figura 34. Pregunta #5 desarrollada en la encuesta específica	98
Figura 35. Pregunta #6 desarrollada en la encuesta específica	99
Figura 36. Pregunta #7 desarrollada en la encuesta específica	100
Figura 37. Esquema servicios laboratorio BD	108
Figura 38. Componentes para el diseño del laboratorio BD	109
Figura 39. Plano de distribución del laboratorio	117
Figura 40. Área de servidores	120
Figura 41. Área de servidores	120
Figura 42. Rampa del área de servidores	121
Figura 43. Piso falso	121
Figura 44. Diagrama refrigeración	122

Figura 45. Vista externa laboratorio BD y área de servidores	123
Figura 46. Vista interna laboratorio BD	127
Figura 47. Vista desde arriba del laboratorio	128
Figura 48. Distribución laboratorio BD	129
Figura 49. Distribución laboratorio BD	129
Figura 50. Distribución laboratorio BD	130
Figura 51. Cotas de emplazamiento para las piernas en puestos de trabajo sentado	132
Figura 52. Altura del plano de trabajo para puesto sentado (cotas en mm)	133
Figura 53. Arco de manipulación vertical en el plano sagital	133
Figura 54. Arco horizontal de alcance del brazo y área de trabajo sobre una mesa	134
Figura 55. Servicios del laboratorio de bases de datos	136
Figura 56. Esquema de servicio de virtualización	139

LISTA DE ANEXOS

	Pág.
Anexo A. Encuesta de carácter general para levantamiento de requerimientos del diseño de laboratorio	148
Anexo B. Encuesta de carácter específico para levantamiento de requerimientos del diseño de laboratorio	152

GLOSARIO

ACADEMIA: Institución oficial constituida por personas destacadas en las letras, las artes o las ciencias, que realizan colectivamente determinadas actividades.

AMAZON: Es una compañía estadounidense de comercio electrónico y servicios de cloud computing a todos los niveles

BACKUP: Es la copia total o parcial de información importante del disco duro, CDs, bases de datos u otro medio de almacenamiento.

CLUSTERING: El término clúster se aplica a los conjuntos o conglomerados de computadoras construidos mediante la utilización de hardwares comunes y que se comportan como si fuesen una única computadora

DASHBOARDS: Es una Interfaz gráfica de usuario que yace tanto en consolas de videojuegos como en algunos sistemas operativos. Es una interfaz donde el usuario puede administrar el equipo y/o software.

DATO: Un dato es un documento, una información o un testimonio que permite llegar al conocimiento de algo o deducir las consecuencias legítimas de un hecho.

DISEÑO: Actividad creativa que tiene por fin proyectar objetos que sean útiles y estéticos

GOOGLE DRIVE: Es un servicio de alojamiento de archivos

GRIDCOMPUTING: La computación grid es una tecnología innovadora que permite utilizar de forma coordinada todo tipo de recursos que no están sujetos a un control centralizado.

HOSTING: Alojamiento web es el servicio que provee a los usuarios de Internet un sistema para poder almacenar información, imágenes, vídeo, o cualquier contenido accesible vía web

INFORMACIÓN: La información está constituida por un grupo de datos ya supervisados y ordenados, que sirven para construir un mensaje basado en un cierto fenómeno o ente. La información permite resolver problemas y tomar decisiones, ya que su aprovechamiento racional es la base del conocimiento.

INTERACTIVO: Que permite una interacción, a modo de diálogo, entre la máquina y el usuario.

LABORATORIO: Un laboratorio es un lugar que se encuentra equipado con los medios necesarios para llevar a cabo experimentos, investigaciones o trabajos de carácter científico o técnico.

MÁQUINA VIRTUAL: Una máquina virtual es un software que emula a una computadora y puede ejecutar programas como si fuese una computadora real.

METODOLOGÍA: Conjunto de métodos que se siguen en una investigación científica, un estudio o una exposición doctrinal

MULTIPLATAFORMA: Que puede utilizarse en diversos entornos o sistemas operativos

OPEN SOURCE: Se define por la licencia que lo acompaña, que garantiza a cualquier persona el derecho de usar, modificar y redistribuir el código libremente.

OUTSOURCING: Es la subcontratación de terceros para hacerse cargo de ciertas actividades complementarias a la actividad principal.

PRÁCTICA DE LABORATORIO: Es el tipo de clase que tiene como objetivos instructivos fundamentales que los estudiantes adquieran las habilidades propias de los métodos de la investigación científica, amplíen, profundicen, consoliden, realicen, y comprueben los fundamentos teóricos de la asignatura mediante la experimentación empleando los medios de enseñanza necesarios, garantizando el trabajo individual en la ejecución de la práctica.

PROCESADOR: Componente electrónico donde se realizan los procesos lógicos

RACK: Un rack es un soporte metálico destinado a alojar equipamiento electrónico, informático y de comunicaciones

REQUERIMIENTO: Un requerimiento es una necesidad documentada sobre el contenido, forma o funcionalidad de un producto o servicio.

SERVIDOR: Un servidor es un nodo que, formando parte de una red, provee servicios a otros nodos denominados clientes

TECNOLOGÍA: Es el conjunto de conocimientos técnicos, científicamente ordenados, que permiten diseñar y crear bienes y servicios que facilitan la adaptación al medio ambiente y satisfacer tanto las necesidades esenciales como los deseos de la humanidad.

WI-FI: Es un mecanismo de conexión de dispositivos electrónicos de forma inalámbrica.

ACRONIMOS

BI: Business Intelligence

CPD: centro de procesamiento de datos

CPU: Central Processing Unit

DB: Data Base

DBMS: Data Base Management System

GISIC: Grupo de Investigación sobre el sostenimiento de la Iglesia católica

HVAC: heating, ventilating, and air conditioning

IP: Internet Protocol

IT: Information Technology

LAN: Local Area Network

LR: Laboratorio remoto

LV: Laboratorio virtual

MAN: Metropolitan Area Network

MIT: Massachusetts Institute of Technology

NFPA: National fire protection association

PC: Personal Computer

Pyme: pequeña y mediana empresa

SQL: Structured Query Language

UPS: uninterruptible power supply

UPS: Uninterruptible Power Supply

WAN: Wide Area Network

XML: Extensible Markup Language

RESUMEN

Este trabajo de grado aborda el diseño el laboratorio de bases de datos para la Universidad Católica de Colombia, en sí, dedicado al programa de Ingeniería de Sistemas. El diseño de este laboratorio hace parte del plan de mejoramiento del programa y permite fortalecer el área curricular de administración de la información. Para llevar a cabo este diseño, se realizó un análisis actual de los laboratorios que apoyan el programa de Ingeniería, su estado actual y su forma de operar, todo esto, con el fin de identificar los nuevos servicios e infraestructura que el nuevo laboratorio debe tener, con el fin de apoyar, de mejor forma, los procesos de enseñanza en el programa.

ABSTRACT

This degree project deals with the design of the laboratory of databases for the Universidad Católica de Colombia, itself, dedicated to the engineering of systems program. The design of this laboratory is part of the program improvement plan and allows to strengthen the curricular area of information management. In order to carry through this design, was a current analysis of laboratories that support the program of engineering, its current state and its form to operate, all this, in order to identify new services and infrastructure that the new lab must have, in order to support, at best, the processes of teaching in the program.

INTRODUCCIÓN

A través de la formación científica, la ingeniería ha sido considerada como la ciencia o la rama que permite el estudio de la tecnología por medio de recursos disponibles, que conllevan a soluciones integrales de problemas específicos que buscan mejorar la calidad de vida de las personas; por esta razón provee fundamentos que aportan en el desarrollo integral del ingeniero a través de la aplicación de prácticas experimentales, que toman como base el conocimiento científico y su relación con la investigación.

De esta manera, diferentes universidades en el mundo, por citar algunas, Massachusetts Institute of Technology (MIT) y la Universidad de Lima, apoyan la formación investigativa, científica y académica del estudiante por medio de distintos laboratorios que como fin, brindan competencias para adquirir conocimientos y desarrollar habilidades teórico-prácticas para dar solución a problemáticas reales. En esencia, los laboratorios en ingeniería “han ocupado tradicionalmente un papel muy importante en las escuelas de ingeniería, pues en ellos, los estudiantes pueden aplicar sus conocimientos teóricos a situaciones prácticas. Si se analiza desde el punto de vista nacional, en Colombia, muchas Universidades han implementado laboratorios ingenieriles en distintas ramas del conocimiento con el fin de apoyar el desarrollo educativo, docente, investigativo y de extensión.

En este sentido, la Universidad Católica de Colombia dada a su gran calidad, desempeño y reconocimiento académico cuenta con laboratorios ingenieriles que son de gran calidad pero deben ser ampliados y mejorados y cada vez más, actualizados para que enfrenten nuevos retos y tendencias tecnológicas del

conocimiento. Por tal motivo y dados los procesos de calidad que atraviesa la Universidad y por ende el programa de Ingeniería de Sistemas, se deben afianzar más sus laboratorios y más, en áreas especializadas que combinan las bondades del plan de estudios. Por lo anterior, este trabajo de grado está orientado al diseño del laboratorio de base de datos que sin duda alguna, es uno de los componentes curriculares de mayor relevancia para el programa y para la articulación del grupo de investigación GISIC. Este diseño incorpora elementos innovadores como virtualización, manejo de grandes cantidades de datos, esquemas Big Data, Clustering, sistemas de información y manejo de sistemas distribuidos. A su vez, responde a las diversas necesidades del plan de estudios asociadas al área curricular denominada Administración de la Información.

1. ANTECEDENTES

Las actividades de laboratorio constituyen un hecho diferencial propio de la enseñanza de la ciencia y la tecnología. “Hace casi trescientos años que John Locke propuso la necesidad de que los estudiantes realizaran trabajo práctico en su educación, y a finales del siglo XIX ya formaba parte integral del currículo de ciencias en Inglaterra y Estados Unidos.”¹

Por otra parte debido a la necesidad de crear sistemas de apoyo al estudiante, para realizar las prácticas del laboratorio. Optimizando tiempo, recursos requeridos y en un ambiente controlado que brinde seguridad, surgen los laboratorios virtuales adaptivos, como un elemento de apoyo para el proceso de aprendizaje de los estudiantes.

Los laboratorios virtuales han ido adentrándose en las prácticas pedagógicas, desde hace más de 25 años aproximadamente, cuando inició su uso. La primera aproximación a los laboratorios virtuales aparece en el año 1984, donde surge el concepto de instrumento virtual y sus características se determinaron de acuerdo a los fundamentos de programación. En 1992 aparece el término laboratorio virtual, para describir la programación orientada a objetos en el desarrollo de un laboratorio de simulación. En 1994 se presenta un estudio realizado por la Universidad de Vanderbilt en EEUU, en el que se desarrolla un laboratorio virtual basado en simulación como apoyo a las prácticas tradicionales y que concluyó, con la necesidad de esta herramienta para aprender las habilidades básicas y el

¹ Barrera, O. Valdez, P. El trabajo práctico en la enseñanza de las ciencias: una revisión [En línea] < http://www.raco.cat/index.php/ensenanza/article/viewFile/21466/93439%3Forigin%3Dpublication_detail > [Citado en 2 de septiembre de 2014]

manejo de los equipos, lo cual optimizaba tanto el tiempo de los alumnos como el del personal de laboratorio.²

En el mundo, universidades como por ejemplo el MIT-Massachusetts Institute of Technology por citar una, posee una serie de laboratorios; *lab computer science*, *lab civil and environmental engineering*, *robotics and artificial intelligence*, por nombrar algunos de los muchos con los que cuenta, en donde la investigación de vanguardia está teniendo lugar.

Sin embargo este trabajo de grado está enfocado hacia los laboratorios en ingeniería de sistemas, más específicamente, laboratorios asociados a las bases de datos. En el mundo existen muchas universidades que tienen este tipo de laboratorios, que satisfacen diferentes servicios como por ejemplo; análisis de datos, control de la información, administración de la información, entre otros y que a su vez apoyan el plan de estudios de los programas de ingeniería. En Colombia por citar una universidad, la universidad de los Andes, posee 2 laboratorios de propósito general en donde se pueden trabajar temas asociados a las bases de datos, la universidad posee licencias ORACLE y un convenio educativo con AMAZON y en donde además los laboratorios están virtualizados es decir los estudiantes pueden acceder a ellos de forma remota y elaborar sus respectivas prácticas e investigaciones.

Por otra parte, dado el desarrollo de este trabajo de grado se precisa mencionar que el programa de ingeniería de sistemas, desde su creación en el año 1986 tiene aproximadamente una comunidad de 2500 egresados y todos ellos de alguna u otra forma han desarrollado prácticas de laboratorio. Sin embargo en el transcurso de estos años no ha existido una articulación clara de lo que son los laboratorios

² Novoa, Norberto; Flórez, Héctor. Los laboratorios virtuales adaptivos y personalizados en la educación superior En: revistavinculos [En línea](septiembre 2011). Disponible en: <<http://revistavinculos.udistrital.edu.co/files/2012/12/Los-laboratorios-virtuales.pdf>> [citado en 14 agosto de 2014].

del programa de ingeniería de sistemas y además de ellos, dado al contexto en el que se desenvuelve el programa, al ser un área tecnología los cambios a nivel curricular se presentan muy rápidamente.

En este sentido, este trabajo de grado se abordara el tema concreto, el diseño del laboratorio de bases de datos que hace parte del área curricular de administración de la información. Por otro lado, Se destaca actualmente que la facultad de ingeniería cuenta con diversos laboratorios, dentro de ellos hay un laboratorio de ingeniera de sistemas que está compuesto por 8 salas, donde cada una de ellas cuenta con un aproximado de 17 equipos de cómputo y algunos servidores que apoyan las prácticas de laboratorio.

2. JUSTIFICACIÓN

Los laboratorios en el ámbito del área de la Ingeniería son determinantes para el avance de la ciencia y la tecnología, de hecho, son los escenarios base para la experimentación dando soporte a los conocimientos teóricos adquiridos. En las Universidades, las Facultades de Ingeniería siguen esta premisa, por ello, los laboratorios toman gran importancia en el currículo de los planes de estudios. Para los programas de ingeniería de sistemas, que están muy cerca de los desarrollos tecnológicos, son imprescindibles para generar competencias y habilidades específicas en los estudiantes.

Por otra parte el crecimiento de la información está llevando a cambiar el mundo, ya que la información debe ser procesada y filtrada, puesto que es el activo más importante de las organizaciones, de hecho, dentro del plan de estudio del programa de ingeniería de sistemas existe el área curricular de administración de la información, donde esto toma gran relevancia, dado que es una de las grandes necesidades del mercado. Es tanto el crecimiento de la información, que las bases de datos han surgido grandes avances, hasta el punto de hablar de bases de datos no estructuradas.

Por su parte en la Universidad Católica de Colombia, el programa de ingeniería de sistemas, en la búsqueda de mejorar y estar a la vanguardia tecnológica de todos estos nuevos retos que se presentan en el mundo, está tomando medidas para continuar en el avance tecnológico, sin embargo, actualmente la universidad no cuenta con un elemento curricular práctico como el laboratorio de base de datos, que permita afianzar los conocimientos asociados a este tipo de temática. Razón por la cual este trabajo de grado busca diseñar el laboratorio de base de datos, con el fin de que el estudiante obtenga conocimiento, adquiera información adecuada y

desarrolle sus habilidades prácticas en distintas áreas de su formación académica, generando así beneficios, algunas como capacitaciones, realización de prácticas, oportunidades de negocio a empresas pymes, sistemas de información, mecanismo de clúster de información, servicios a terceros, generando recursos externos que aportaran en el desarrollo de la universidad con miras a un crecimiento tecnológico que contribuyan en las competencias y habilidades ingenieriles del estudiante, satisfaciendo y cubriendo las necesidades del plan de estudios.

3. PLANTEAMIENTO DEL PROBLEMA

Desde siempre existe una constante preocupación por el nivel educativo de las naciones y más, cuando se focaliza hacia la educación superior donde las Universidades intensifican los avances en ciencia y tecnología. Sin embargo, lo anterior es inalcanzable si no se tiene como base una gran fundamentación práctica que por supuesto, es alcanzada gracias al uso de laboratorios en las áreas de ingeniería.

En la Universidad Católica de Colombia, en el Programa de Ingeniería de Sistemas lo referente a los laboratorios que apoyan el plan de estudios, son de gran relevancia y más si se analiza desde el punto de vista curricular. Sin embargo, dado el ámbito tecnológico donde se mueve el programa y a su vez el plan de estudios, evidencian que la infraestructura actual de los laboratorios debe ser mejorada para que responda a nuevos enfoques tecnológicos para el área curricular asociada a la administración de la información. Además, se precisa resaltar que actualmente el programa requiere mayor articulación entre las asignaturas y los laboratorios con el fin de que los conocimientos teórico-prácticos puedan desarrollarse de mejor forma lo que generará, directamente, una aportación mayor a los estudiantes en la construcción de sus competencias ingenieriles y permitirá articular el componente investigativo del programa asociado al grupo GISIC. Por ello, se destaca que los laboratorios asociados al programa requieren mayor atención por parte de las directivas del programa, por ello, el plan de mejoramiento del programa contempla, entre otros, el diseño del laboratorio de bases de datos para que aporte debidamente al currículo del programa en áreas asociadas a virtualización, manejo de grandes cantidades de datos, esquemas Big Data, Clustering, sistemas de información, y manejo de sistemas distribuidos, áreas

que actualmente se contemplan en un enfoque mínimo en el programa y más con respecto a los laboratorios.

Con base en lo anterior se plantea la pregunta, ¿Qué características tanto de hardware como software, son las requeridas para dar soporte al laboratorio de base de datos para el programa de ingeniería de sistemas de la Universidad Católica de Colombia, con el fin de responder a los nuevos enfoques tecnológicos y de mercado que demanda el mundo?

4. OBJETIVOS

4.1 OBJETIVO GENERAL

- Diseñar un laboratorio de bases de datos para la Universidad Católica de Colombia

4.2 OBJETIVOS ESPECÍFICOS

- Realizar el levantamiento del estado actual del laboratorio de base de datos.
- Realizar el levantamiento de requerimiento del laboratorio de base de datos.
- Diseñar el laboratorio de bases de datos.
- Validar el diseño propuesto del laboratorio de bases de datos.

5. HIPÓTESIS

Si se logra diseñar el laboratorio de bases de datos, adscrito al área curricular de la administración de la información, será posible aumentar las competencias ingenieriles del futuro egresado y así mismo apoyar de forma directa las líneas de investigación del programa de ingeniería de sistemas de la Universidad Católica de Colombia.

6. MARCO CONCEPTUAL

En este ítem se definirán algunos conceptos de gran importancia para la realización de este proyecto.

6.1 INFRAESTRUCTURA

“Conjunto de todos los elementos tecnológicos hardware y software: computadores, portátiles, impresoras, etc.”³ Que pueden integrar un proyecto o respaldar una operación.

6.2 HARDWARE

“Es la parte física de cualquier dispositivo electrónico o informático, es usual que sea utilizado en una forma más amplia, generalmente para describir componentes físicos de una tecnología, incluyendo equipos de cómputo, periféricos, redes, cableado y cualquier otro elemento físico involucrado.”⁴

6.3 SOFTWARE

“Es un término informático que hace referencia a un programa o conjunto de programas de cómputo que incluye datos, procedimientos y pautas y que permite realizar distintas tareas en un sistema informático.”⁵

6.4 REDES

Es un sistema de comunicaciones capaz de facilitar el intercambio de datos informáticos entre grupos de computadoras y usuarios.

³ administración de la infraestructura tecnológica. [En línea] (Agosto 2011). Disponible En: < http://quindio.gov.co/home/docs/items/item_100/P-SAD-71Administraciondelainfraestructuratecnologica.pdf> [citado en 08 octubre de 2014].p.1

⁴ Ibid.,p.2

⁵ Significado del hardware. [En línea]. Disponible En:< <http://www.significados.info/software/>> [citado en 3 de septiembre de 2014]

6.4.1 Clasificación de las redes

- **LAN:** son redes limitadas en espacio, en si para cortas distancias y sirven para entornos tales como: oficinas y centros de cómputo de un pequeño tamaño. De igual forma, son redes veloces y son utilizadas tanto alámbrica como inalámbrica.
- **MAN:** es una red que brinda cobertura en un área geográfica mucho más amplia, permitiendo conectar varias LAN a altas velocidades.
- **WAN:** esta red se expande en ciudades, países, continentes, etc. Pero a diferencia de la LAN no poseen la misma velocidad, aunque si tiene la capacidad de transportar un mayor volumen de datos.

6.5 VIRTUALIZACIÓN

“La virtualización es la tecnología que permite a las empresas consolidar su infraestructura de TI para poder así a lograr un mayor aprovechamiento de los recursos. Es una capa abstracta que desacopla el hardware físico del sistema operativo para brindar una mayor flexibilidad y utilización de los recursos de TI.”⁶

6.6 ACCESO REMOTO

Es un medio de acceso mediante una computadora a un recurso físico ubicado en otra computadora por medio de una red local o externa.

⁶ Virtualización de Servidores [En línea]. Disponible En:
<http://www.usmp.edu.pe/vision2012_lima/SEMINARIOS/seminarios/Virtualizacion_de_Servidores.pdf>
[citado en 9 de octubre de 2014]

6.7 LABORATORIO

Son empleados como un elemento de enseñanza y aprendizaje, creando así un lazo entre la teoría y la práctica. Una definición cercana podría ser:

Un lugar equipado con diversos instrumentos de medición, entre otros, donde se realizan experimentos o investigaciones diversas, según la rama de la ciencia a la que se enfoque. Dichos espacios se utilizan tanto en el ámbito académico como en la industria y responden a múltiples propósitos, de acuerdo con su uso y resultados finales, sea para la enseñanza, para la investigación o para la certificación de la industria⁷.

6.8 LABORATORIOS VIRTUALES (LV)

Es un sistema computacional accesible vía Internet, mediante un simple navegador, se puede simular un laboratorio convencional en donde los experimentos se llevan a cabo siguiendo un procedimiento similar al que se sigue en un laboratorio convencional, pudiendo inclusive ofrecer la visualización de instrumentos y fenómenos mediante objetos dinámicos, incluyendo imágenes y animaciones.⁸

6.9 LABORATORIO REMOTO (LR)

Se pueden considerar como una evolución de los LV. En este caso al sistema computacional se les agregan instrumentación, control y acceso a equipos de laboratorio reales. Ya no hablamos de llevar a cabo prácticas en un simulador, sino que se trata de realizar actividades prácticas de forma local o remota a través de una Intranet o Internet, permitiendo la transferencia de información entre un proceso real y los estudiantes de manera unidireccional o bidireccional. Bajo este esquema el

⁷ Ingeniería: La importancia de los laboratorios. [En línea] (Diciembre 2006). Disponible En: <<http://www.imcyc.com/revistact06/dic06/INGENIERIA.pdf>> [citado en 30 julio de 2014].

⁸ Lorandi, A. Hermida, H. Hernández, J. Duran, H. Los Laboratorios Virtuales y Laboratorios Remotos en la Enseñanza de la Ingeniería. En: *academiajournals* [en línea]. Vol 4 (2011) <<http://academiajournals.com/downloads/LorandiLabsEd11.pdf>> [citado en 07 de Octubre de 2014].p.3

estudiante utiliza y controla los recursos disponibles en un laboratorio, mediante el uso de tarjetas de adquisición de datos, sensores e instrumentos de medida con interfaces de red y software.⁹

6.10 CABLEADO ESTRUCTURADO

“Cableado Estructurado es el cableado de un edificio o una serie de edificios que permite interconectar equipos activos, de diferentes o igual tecnología permitiendo la integración de los diferentes servicios que dependen del tendido de cables como datos, telefonía, control, etc.”¹⁰

6.11 RACK

“Es un bastidor destinado a alojar equipamiento electrónico, informático y de comunicaciones. Sus medidas están normalizadas para que sea compatible con equipamiento de cualquier fabricante.”¹¹

6.12 BASES DE DATOS

Una colección de datos interrelacionados almacenados conjuntamente en uno o más ficheros de computadora. “Es sumamente importante y beneficioso para una empresa contar con una base de datos unificada que centralice toda la información y permita a todos sus empleados gestionar dicha base de datos con los permisos y accesos adecuados.”¹²

⁹ Ibid.,p.4

¹⁰ cableado estructurado. [En línea]. Disponible En:< http://materias.fi.uba.ar/6679/apuntes/CABLEADO_ESTRUC.pdf/> [citado en 9 de octubre de 2014]

¹¹ Concepto de rack. [En línea]. Disponible En:< http://www.rigidosc.com.ar/?page_id=20> [citado el 18 de octubre de 2014]

¹² Sicilia, Miguel. Las bases de datos son información almacenada en ficheros [En línea]. Disponible En:< <http://cnx.org/content/m17423/latest/>> [citado en 30 de julio de 2014].

6.13 BASE DE DATOS ESTRUCTURADO

“Las bases de datos no estructurados son repositorios en los que se consolida y ordena esta información, garantizando su integridad y poniéndola a la disposición de toda la organización”¹³.

6.14 BASE DE DATOS NO ESTRUCTURADOS

“El contar con bases de datos eficientes es un componente vital de la infraestructura tecnológica de cualquier organización porque garantiza un excelente tiempo de respuesta a los usuarios y permite contar con un esquema de alta disponibilidad para las aplicaciones empresariales”¹⁴.

6.15 GESTOR DE BASES DE DATOS (SGBD)

“También llamado DBMS (Data Base Management System) como una colección de datos relacionados entre sí, estructurados y organizados, y un conjunto de programas que acceden y gestionan esos datos. La colección de esos datos se denomina Base de Datos o BD, (DB Data Base)”¹⁵.

6.16 DATAMINIG

Es el conjunto de técnicas y tecnologías que permiten explorar grandes bases de datos, de manera automática o semiautomática, con el objetivo de encontrar patrones repetitivos, tendencias o reglas que expliquen el comportamiento de los datos en un determinado contexto. Básicamente, el datamining surge para intentar ayudar a comprender el contenido de un repositorio de datos. Con este fin, hace

¹³ Las bases de datos estructurados [En Línea]. Disponible En:< http://www.ids-ac.com/index.php?option=com_content&view=article&id=20&Itemid=52> [citado en 03 de septiembre de 2014].p.5

¹⁴ Sistemas gestores bases de datos [En línea]. Disponible En:< <http://www.mcgraw-hill.es/bcv/guide/capitulo/8448148797.pdf> > [citado en 30 julio de 2014]

uso de prácticas estadísticas y, en algunos casos, de algoritmos de búsqueda próximos a la Inteligencia Artificial y a las redes neuronales¹⁶.

6.17 BIG DATA

Tendencia en el avance de la tecnología que ha abierto las puertas hacia un nuevo enfoque de entendimiento y toma de decisiones, la cual es utilizada para describir enormes cantidades de datos (estructurados, no estructurados y semi estructurados) que tomaría demasiado tiempo y sería muy costoso cargarlos a una base de datos relacional para su análisis. De tal manera que, el concepto de Big Data aplica para toda aquella información que no puede ser procesada o analizada utilizando procesos o herramientas tradicionales¹⁷.

6.18 MAP REDUCE

El término MapReduce en realidad se refiere a dos tareas separadas y distintas que los programas de Hadoop realizan. El primero es el trabajo de un mapa, que tiene un conjunto de datos y lo convierte en otro conjunto de datos, donde los elementos individuales se desglosan en tuplas (pares clave / valor). El trabajo de reducir toma la salida de un mapa como entrada y combina las tuplas de datos en un conjunto más pequeño de tuplas.¹⁸

6.19 DATA WAREHOUSE

Según Bill Inmon: “Una bodega de datos es una colección de datos, integrados, no volátiles, variante en el tiempo y orientados a temas, organizados para soportar necesidades empresariales”, especialmente soporta los procesos de toma de decisiones gerenciales. La bodega de datos se alimenta de las bases y fuentes de datos de los sistemas operacionales y transaccionales, a través de

¹⁶ Business intelligence- datamining (Minería de datos) [En Línea]. Disponible En:< <http://www.mcgraw-hill.es/bcv/guide/capitulo/8448148797.pdf>> [citado en 14 de agosto de 2014]

¹⁷ ¿Qué es Big Data? [En línea]. Disponible En:< <http://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/>> [Citado en 14 de agosto de 2014]

¹⁸ ¿Qué es MapReduce? [En línea]. Disponible En: < <http://www-01.ibm.com/software/data/infosphere/hadoop/mapreduce/>> [Citado en 9 de octubre de 2014]

un proceso de integración de datos donde los datos extraídos se someten a transformaciones para eliminar las inconsistencias y resumir la información.¹⁹

6.20 HERRAMIENTAS REQUERIDAS EN EL LABORATORIO DE BASES DE DATOS

6.20.1 Sistema gestor de bases de datos

6.20.1.1 MySQL.

Es un sistema gestor de bases de datos relacionales rápido, sólido y flexible. Es idóneo para la creación de bases de datos con acceso desde páginas web dinámicas, así como para la creación de cualquier otra solución que implique el almacenamiento de datos, posibilitando realizar múltiples y rápidas consultas. Está desarrollado en C y C++, facilitando su integración en otras aplicaciones desarrolladas también en esos lenguajes. Es un sistema cliente/servidor, por lo que permite trabajar como servidor multiusuario y de subprocesamiento múltiple, o sea, cada vez que se crea una conexión con el servidor, el programa servidor establece un proceso para manejar la solicitud del cliente, controlando así el acceso simultáneo de un gran número de usuarios a los datos y asegurando el acceso a usuarios autorizados solamente. Es uno de los sistemas gestores de bases de datos más utilizado en la actualidad, utilizado por grandes corporaciones como Yahoo! Finance, Google, Motorola, entre otras.²⁰

¹⁹ Construcción de bodegas de datos [En línea] Disponible En:< <http://bianalytics.biz/index.php/bi-analytics-inteligencia-de-negocios-bi-services/72-construccion-bodegas-datos>> [citado en 14 de agosto de 2014]

²⁰ EcuRed.Sistema gestor de bases de datos, [En línea]. http://www.ecured.cu/index.php/Sistema_Gestor_de_Base_de_Datos. [Citado 29 Octubre 2014].p.7

6.20.1.2 PostgreSQL.

Es un Sistema Gestor de Bases de Datos Relacionales Orientadas a Objetos, derivado de Postgres, desarrollado en la Universidad de California, en el Departamento de Ciencias de la Computación de Berkeley. Es un gestor de bases de datos de código abierto, brinda un control de concurrencia multi-versión (MVCC por sus siglas en inglés) que permite trabajar con grandes volúmenes de datos; soporta gran parte de la sintaxis SQL y cuenta con un extenso grupo de enlaces con lenguajes de programación.²¹

6.20.1 Modelado de datos

6.20.1.1 DBDesigner4. Es un sistema de diseño de base de datos visual que integra el diseño de bases de datos, modelado, creación y mantenimiento en un único entorno, sin fisuras. Combina características profesionales y una interfaz de usuario clara y sencilla de ofrecer la forma más eficiente de manejar sus bases de datos. Es desarrollado y optimizado para el código abierto MySQL-Base de datos para apoyar a los usuarios de MySQL con una herramienta de diseño potente y gratuito. Todas las características específicas de MySQL se han construido para ofrecer la forma más conveniente para diseñar y mantener el control de sus bases de datos MySQL.

6.20.1.2 Power Architect. La herramienta de modelado de datos SQL Arquitecto energía fue creado por diseñadores de almacenamiento de datos y tiene muchas características únicas orientadas específicamente para el arquitecto de almacenamiento de datos. Permite a los usuarios realizar ingeniería inversa de bases de datos existentes, realizan perfiles de datos en bases de datos fuente, y auto-generan metadatos de ETL.

²¹Ibid.,p.8

6.20.1.3 Wwsqldesigner. Es una herramienta de modelado de SQL basado en la web se distribuye bajo la licencia BSD Nueva. Es creado y mantenido por Ondrej Zara en Códigos de Google. WWW SQL Designer le permite dibujar y crear esquemas de bases de datos (diagramas ER) directamente en tu navegador, sin necesidad de ningún programa externo como Flash. Sólo es necesario activar JavaScript. El Diseñador funciona a la perfección en de Mozilla (Firefox, Seamonkey), Exploradores de Internet (6, 7, 8), Safari y óperas. Konqueror funciona, pero la experiencia es limitada.

Muchas de las funciones de base de datos son compatibles, tales como llaves, claves foráneas, comentarios e índices. Puede guardar su diseño (para mayor carga y modificaciones), imprimirlo o exportar como secuencia de comandos SQL. También es posible recuperar (importación) esquema de una base de datos existente.

6.20.1.4 MicroOLAP. Permite para construir su base de datos en entorno visual fácil de entender por el trabajo con la representación gráfica de tablas, columnas, relaciones y así sucesivamente.

Este producto permite extraer tablas, atributos, relaciones, índices y otros objetos de las bases de datos existentes, incluyendo MySQL, Microsoft Access y Sybase ASE / ASA, Oracle, Informix, MSSQL, DB2, DBF y muchos otros que puede acceder a través de OLEDB o ODBC e importación (ingeniería inversa) ellos en su diagrama. Finalmente obtendrá física Entidad Relación Model (Modelo ER) que se puede exportar en secuencia de comandos SQL, o directamente construyó sobre el trabajo del servidor de base de datos MySQL. Después de que su base de datos model / diagrama se modifica, puede sincronizar la base de datos MySQL existente con él.

6.20.1.5 Enterprise Architect. Enterprise Architect es una herramienta comprensible de diseño y análisis UML, cubriendo el desarrollo de software desde el paso de los requerimientos a través de las etapas del análisis, modelos de diseño, pruebas y mantenimiento. EA es una herramienta multi-usuario, basada en Windows, diseñada para ayudar a construir software robusto y fácil de mantener. Ofrece salida de documentación flexible y de alta calidad.

6.20.2 Big data. El software nombrado a continuación, fue seleccionado con base al posicionamiento que tienen en el mercado, por ser software open source puesto que tienen un fácil acceso y porque aportan en las prácticas del laboratorio de bases de datos.

6.20.2.1 Avro.

Es un sistema para la serialización de datos y uno de los principales métodos para transportar información entre las herramientas y aplicaciones Hadoop. Entre otras características, Avro provee a las aplicaciones que la utilizan de un formato de serialización binario, compacto y de rápido transporte siendo compatible con estructuras de datos complejas; un método para escribir datos de forma persistente a través de ficheros; y la implementación del protocolo RPC (Remote Procedure Call), para la comunicación entre ordenadores con una arquitectura maestro-esclavo. Las librerías que ofrece Avro son de uso muy sencillo ya que no requieren de generación de código ya sea para leer o escribir datos o para usar una comunicación de tipo RPC. El funcionamiento de Avro es muy sencillo, al serializar datos Avro crea un esquema JSON (Figura 1) que lo acompañará durante todo el “trayecto”. De esta manera cuando estos datos sean leídos, la aplicación destino tendrá toda la información necesaria para interpretar los datos que le llegan. Esto hace a Avro totalmente compatible con lenguajes dinámicos.²²

²² Serrat., Robert. Big-Data- Análisis de herramientas y soluciones. Barcelona, 2013. Trabajo de grado (Ingeniero Informática).Universidad Politécnica de Cataluña. Facultad Informática. Disponible en: < <http://upcommons.upc.edu/pfc/bitstream/2099.1/19855/1/90807.pdf> >.p.9

Figura 1. Ejemplo de un esquema json

```
{
  "namespace": "ejemplo.avro",
  "type": "record",
  "name": "Usuario",
  "fields": [
    {
      "name": "nombre", "type": "string"
    },
    {
      "name": "numero_favorito", "type": ["int", "null"]
    },
    {
      "name": "color_favorito", "type": ["string", "null"]
    }
  ]
}
```

Fuente: Serrat., Robert. Big-Data- Análisis de herramientas y soluciones. Barcelona, 2013. Trabajo de grado (Ingeniero Informática).Universidad Politécnica de Cataluña. Facultad Informática. Disponible en: <http://upcommons.upc.edu/pfc/bitstream/2099.1/19855/1/90807.pdf> <

6.20.2.2 Zookeeper.

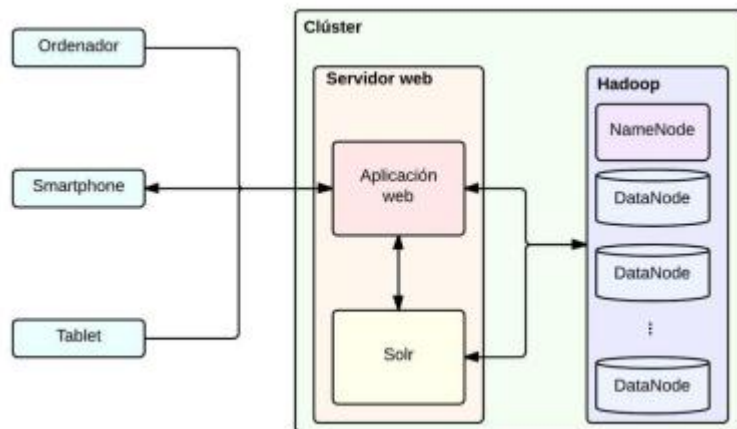
Es un servicio centralizado que se encarga de administrar y gestionar la coordinación entre procesos en sistemas distribuidos. El objetivo de Apache con ZooKeeper es el de librar a los desarrolladores la tarea de implementar funciones de mantenimiento entre sus procesos, como la sincronización entre estos, y ofrecer alta disponibilidad a través de servicios redundantes. ZooKeeper permite la coordinación de procesos a través de un espacio de nombres compartido de estructura jerárquica, parecido a un sistema de ficheros, donde cada nodo recibe el nombre de znode. Los znodes son muy parecidos a los de un sistema de ficheros normal y corriente, cuentan con un path o camino con los elementos separados por la barra "/" y cada uno tiene un padre que es el znode anterior - salvo la raíz "/" que es su propio padre-. Los znodes tampoco pueden ser eliminados si tienen hijos, tal y como pasa con los sistemas de ficheros comunes.²³

²³ Ibid. p.10.

6.20.2.3 Solr.

Apache Solr es un motor de búsqueda basado en el Apache Lucene, escrito en Java y que facilita a los programadores el desarrollo de aplicaciones de búsqueda. Lucene ofrece indexación de información, tecnologías para la búsqueda así como corrección ortográfica, resaltado y análisis de información, entre otras muchas características. Una arquitectura típica de Solr sería la de la Figura 2. Se cuenta con un servidor web, para que los usuarios puedan interactuar y realizar distintos tipos de búsquedas -páginas de tiendas en línea, por ejemplo-, con conexión directa con Solr y que consulta datos mediante este en Hadoop. El servidor web puede ser una de las máquinas con servicios de Hadoop ejecutándose en él o un servidor exclusivamente dedicado a servir peticiones de los usuarios.²⁴

Figura 2. Arquitectura de solr trabajando con hadoop.



Fuente: Serrat., Robert. Big-Data- Análisis de herramientas y soluciones. Barcelona, 2013. Trabajo de grado (Ingeniero Informática).Universidad Politécnica de Cataluña. Facultad Informática. Disponible en: <http://upcommons.upc.edu/pfc/bitstream/2099.1/19855/1/90807.pdf> >.

²⁴ Ibid. p.11.

6.20.2.4 Chukwa.

Es una herramienta principalmente pensada para trabajar sobre logs y realizar análisis. Está construido por encima de HDFS y MapReduce, por lo que hereda su escalabilidad y robustez. La principal motivación a la hora de desarrollar Chukwa fue precisamente que Hadoop no termina de trabajar bien con sistemas de logs ya que está más optimizado para trabajar con pocos ficheros de mayor tamaño; a contraposición de los sistemas de logs, que son directorios con un gran número de ficheros pequeños. Para cumplir con este propósito Chukwa ofrece un sistema flexible para recolectar datos de forma distribuida y para realizar su procesamiento que, a la vez, es adaptable a las nuevas tecnologías de almacenamiento que van apareciendo.²⁵

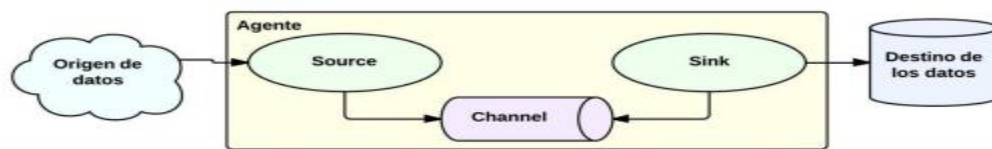
6.20.2.5 Flume.

Es una herramienta distribuida para la recolección, agregación y transmisión de grandes volúmenes de datos. Ofrece una arquitectura basada en la transmisión de datos por streaming altamente flexible y configurable pero a la vez simple. Al tener un origen de datos configurable, se adapta prácticamente a cualquier tipo de situación: monitorización de logs, descarga de información de redes sociales o mensajes de correo electrónico, entre muchas otras. Los destinos de los datos también son altamente configurables de manera que el uso de Flume no va ligado exclusivamente con HDFS o incluso Hadoop. La arquitectura de Flume está basada en agentes, que son procesos encargados de recolectar datos y enviarlos a su destino. A su vez, el flujo de datos viene determinado por una unidad llamada evento que, como pasaba con los chunks de Chukwa, está formado por datos y metadatos.²⁶

²⁵ Ibid. p.12.

²⁶ Ibid., p.13.

Figura 3. Componentes de un agente de flume



Fuente: Serrat., Robert. Big-Data- Análisis de herramientas y soluciones. Barcelona, 2013. Trabajo de grado (Ingeniero Informática).Universidad Politécnica de Cataluña. Facultad Informática. Disponible en: <http://upcommons.upc.edu/pfc/bitstream/2099.1/19855/1/90807.pdf> <

6.20.2.6 Cassandra.

Cassandra es una base de datos NoSQL, mayormente desarrollada por Datastax aunque empezó como un proyecto de Facebook, escrita en Java y open-source - también es un proyecto Apache-. Entre sus características se encuentra la de ofrecer alta disponibilidad de los datos junto con una gran capacidad para escalar linealmente, además de ser tolerante a fallos -no tiene un punto de fallo único- y compatible con hardware de bajo presupuesto o infraestructuras en la nube.²⁷

6.20.2.7 Oozie.

Es un planificador de workflows para sistemas que realizan trabajos o procesos Hadoop. Proporciona una interfaz de alto nivel para el usuario no técnico o no experto y que gracias a su abstracción permite a estos usuarios realizar flujos de trabajo complejos. Oozie funciona como un motor de workflows a modo de servicio que permite lanzar, parar, suspender, retomar y volver a ejecutar una serie de trabajos Hadoop -no solamente MapReduce- basándose en ciertos criterios, como temporales o de disponibilidad de datos. Los flujos de trabajo Oozie son grafos no cíclicos directos -también conocidos como DAGs- donde cada nodo es un trabajo o

²⁷ Ibid., p.14.

acción con control de dependencia, es decir, que una acción no puede ejecutarse a menos que la anterior haya terminado.²⁸

6.20.3 Data wharehouse

6.20.3.1 Hive.

Facilita la creación, consulta y administración de grandes volúmenes de datos distribuidos en forma de tablas relacionales. Cuenta con un lenguaje derivado de SQL, llamado Hive QL, que permite realizar las consultas sobre los datos. Por esta misma razón, se dice que Hive lleva las bases de datos relacionales a Hadoop. A su vez, Hive QL está construido sobre MapReduce, de manera que se aprovecha de las características de éste para trabajar con grandes cantidades de datos almacenados en Hadoop. Esto también provoca que Hive no ofrezca respuestas en tiempo real.²⁹

6.20.3.2 Pig.

Es una herramienta para analizar grandes volúmenes de datos mediante un lenguaje de alto nivel -PigLatin- que está diseñado para la paralelización del trabajo. Mediante un compilador se traducen las sentencias en PigLatin a trabajos MapReduce sin que el usuario tenga que pasar a programar ni tener conocimientos sobre ellos. Así pues PigLatin es un lenguaje fácil de programar ya que se trata de un lenguaje textual y convierte las paralelizaciones en dataflows o flujos de datos, conceptos más sencillos para los usuarios no expertos o sin conocimiento en el mundo de la paralelización. Además también se encarga de optimizar de manera automática los programas escritos por el usuario, respetando la coherencia de los datos, antes de traducirlos a trabajos MapReduce. Adicionalmente, también permite la creación de funciones por parte del usuario para realizar procesamientos de propósito especial y que no se incluya en las operaciones básicas de PigLatin.³⁰

²⁸ Ibid., p.15.

²⁹ Ibid., p.16.

³⁰ Ibid., p.17.

6.20.3.3 Oracle Business Intelligence Suite.

Es una solución integrada de productos de inteligencia empresarial (BI) con cuadros de mando, un completo sistema de consultas ad hoc, alertas e información proactiva, informes financieros y corporativos, datos predictivos en tiempo real y análisis desconectado entre otras funciones. Los componentes de Oracle Business Intelligence Suite se basan en una arquitectura web orientada a servicios que se integra en la infraestructura empresarial existente, con el menor coste total de propiedad. Oracle Business Intelligence Suite Enterprise Edition Plus proporciona el conocimiento completo y relevante al alcance de cada una de las personas de la organización, simplifica la toma de decisiones y mejora la productividad de los procesos empresariales. Con el Business Intelligence puede convertir los datos almacenados de su empresa en información útil de la situación actual y de la evolución de su empresa. Puede establecer diferentes fuentes de datos, ya sea de su ERP, CRM, una base de datos, ficheros en Excel de presupuestos, ficheros de texto..., y analizar toda la información conjuntamente.³¹

6.20.3.4 Microsoft Business Intelligence.

Microsoft Business Intelligence (BI), es un conjunto completo de productos de cliente, servidor y software formado por Microsoft Office System 2007, SQL Server 2005 y Microsoft Performance Point, que están totalmente integrados con el fin de admitir todos los aspectos de la toma de decisiones. Gracias a esta integración y a las posibilidades probadas de la plataforma de SQL Server 2005, puede confiar en que Microsoft Business Intelligence proporciona la información importante o crucial a toda la organización. Realiza el seguimiento y el análisis de la información financiera, operacional, de clientes y de recursos humanos de la organización proporcionando a los usuarios la información correcta en el momento adecuado en un formato fácil de usar. El uso de tarjetas de puntuación personalizadas basadas

³¹ Softwareseleccion. Oracle Business Intelligence Suite. [En línea].<
<http://www.softwareseleccion.com/oracle+business+intelligence+suite-p-267>> [Citado en 29 de Octubre de 2014]

en Web amplía el alcance de la inteligencia empresarial y permite que todos los empleados participen en el proceso de administración de rendimiento. Las funciones de análisis avanzado y visualización de datos permiten a los usuarios empresariales tomar decisiones mejor fundadas y con mayor rapidez. Microsoft Business Intelligence permite planificar, presupuestar y realizar pronósticos de un modo más eficaz ayudando a los usuarios a crear modelos detallados y planes flexibles y sincronizados entre departamentos y jerarquías organizativas. Estas soluciones están totalmente integradas con Microsoft Dynamics AX, NAV y CRM.³²

6.20.4 Data mining

6.20.4.1 Pentaho BI. Es un software de Business Intelligence (BI) ofrecido como solución libre por parte de Eclipse. Es un software desarrollado para poder realizar ciertas tareas como: Reporting, Análisis, Indicadores (Dashboards), Minería de datos (Data Mining) y Workflow de operaciones. El sistema Pentaho BI ayuda a las empresas a trabajar de forma más eficiente y efectiva. Se trata de la herramienta Open Source líder para Business Intelligence. Al estar basado en Java, la suite Pentaho BI es multiplataforma (Linux, Windows, etc.) y sólo se necesita un navegador para poder acceder a todas sus funcionalidades.

6.20.4.2 Weka.

(Waikato Environment for Knowledge Analysis) Maneja lenguaje Java, es una de las herramientas para aplicación de tareas de data mining más reconocidas, que permite proceso previo, clustering o generación de grupos de datos, clasificación, regresiones, visualización y selección de propiedades. Su técnica se basa en la hipótesis que los datos se encuentran disponibles en un solo archivo o

³² Softwareseleccion. Microsoft Business Intelligence. [En línea].<
<http://www.softwareseleccion.com/microsoft+business+intelligence-p-1372>> [Citado en 29 de Octubre de 2014]

relacionados, donde cada dato está ubicado de acuerdo a su atributo, por ejemplo cliente, edad, género, etc.³³

6.20.4.3 Rapid Miner.

Es un entorno para aprendizaje mecánico y experimentos de data mining, utilizado tanto en investigación como en tareas de día a día por diferentes empresas. Produce sus resultados en archivos XML y cuentan con la interface gráfica del mismo programa. Proveen más de 500 operadores para los principales procesos de aprendizaje en máquina y al tiempo combina esquemas y atributos de evaluación.³⁴

6.20.4.4 Orange. “Es un componente de minería de datos y también es un software de aprendizaje de máquina, que permite una programación visual, rápida y versátil para un análisis exploratorio de datos, aunque es una herramienta poderosa, sigue siendo amigable e intuitiva. También permite pre-procesamiento, filtros de información, modelación de datos, evaluación de modelos y técnicas de exploración.”³⁵

6.20.4.5 JHepWork.

Diseñado para científicos, ingenieros y estudiantes, es una herramienta gratuita y de uso libre, que permite el análisis de datos mediante la creación de un entorno comprensible, amigable y adaptable a programas comerciales. Contiene librerías científicas en Java para funciones matemáticas, y algoritmos de minería de datos.

³³ Ortiz, Paola. 6 Herramientas gratuitas para analisis de datos. [En linea].< <http://mprende.co/gesti%C3%B3n/6-herramientas-gratuitas-para-an%C3%A1lisis-de-datos>> [Citado en 29 de Octubre de 2014],p.22.

³⁴ Ibid., p.23.

³⁵ Ibid., p.24.

Esta herramienta es un poco más avanzada y se requiere más alto conocimiento, el lenguaje usado es Jython, aunque también funciona a la perfección en Java.³⁶

6.20.4.6 Data Mining Add-In for Excel. “Es una herramienta que se puede instalar a la suite de Office, que incluye herramientas de análisis de tablas y una función adicional llamada Data Mining Client, estas herramientas serán vistas al detalle en la próxima entrada.”³⁷

6.20.5 Virtualización. En este apartado, explicaremos acerca de la tecnología de virtualización y un énfasis del porque utilizarla como herramienta dentro de las instalaciones. Para ello describiremos en qué consiste y que puede aportar en las prácticas.

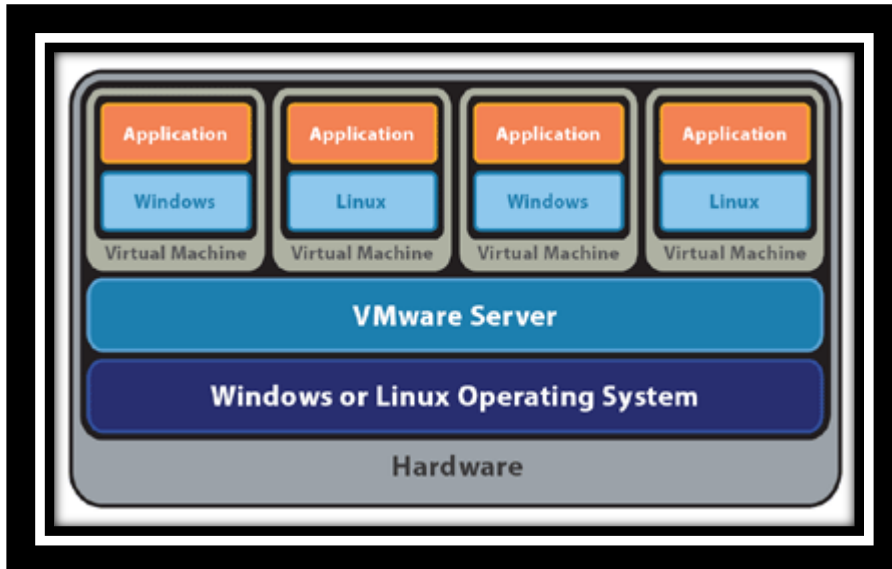
La virtualización es una técnica que posibilita la ejecución de una o más máquinas virtuales sobre una única máquina física. Cada máquina virtual tiene asignados, de forma independiente al resto, un conjunto de recursos de hardware (procesador, memoria, almacenamiento, dispositivos de E/S) y ejecuta su propia copia del sistema operativo (Linux, Solaris, *BSD, Windows). El software de virtualización planifica la ejecución de las máquinas virtuales y gestiona el acceso compartido a los recursos del hardware real y disponible.³⁸

³⁶ Ibid., p.25.

³⁷ Ibid., p.26.

³⁸ Jorge Maldonado Mahauad. DISEÑO DE UN CENTRO DE DATOS BASADO EN ESTANDARES CASO PRÁCTICO: DISEÑO DEL CENTRO DE DATOS DEL COLEGIO LATINOAMERICANO. Cuenca, 2010. Trabajo de grado (Ingeniero en Sistemas). Universidad de Cuenca. Facultad de Ingeniería

Figura 4. Servidor con virtualización



Fuente: < <http://www.consultges.com/servicios.html>>

La virtualización permite crear y alojar una variedad de servidores en una misma computadora física. Esto permite optimizar y reducir el uso de recursos como CPU, memoria, disco duro y entre otras. Una facilidad de la virtualización es la realización de backup y configuración de un servidor virtual, logrando estar en un único archivo.

Además incrementa la seguridad, ya que puede utilizar servidores aislados para utilizarlos en tareas diferentes. Existen 2 tipos de virtualización: Plataforma y de recursos.

6.20.5.1 Virtualización de plataforma. Consiste en simular máquinas virtuales para poder ejecutar varios sistemas operativos en una misma máquina.

Las máquinas virtuales creen que los recursos de los que disponen les pertenecen, y ven a otras máquinas virtuales como sistemas totalmente independientes. De esta forma, los servidores que conocíamos hasta hoy pasan a ser alojados en entes lógicos (máquinas virtuales) destinados a convivir con otros y a compartir los recursos físicos de los que disponen; cuando hablamos de consolidación de

servidores hablamos de virtualización de plataforma. Dependiendo de cómo sea gestionada esta compartición y la convivencia de las máquinas virtuales, y cómo éstas sean organizadas e integradas dentro de una determinada infraestructura virtual estaremos hablando de un tipo u otro de virtualización de plataforma.³⁹

6.20.5.2 Virtualización de recursos. Consiste en simular recursos, un ejemplo es combinar componentes pequeños para simular un recurso mayor o dividir un recurso grande en varios más pequeños virtuales. Un ejemplo claro es:

- **Clúster**

Combinan múltiples computadoras en una gran meta computadora. Los cluster son supercomputadores de bajo costo. Un grupo de PCs conformados de hardware barato conectados en red que se comportan como un solo ordenador de gran potencia.

El laboratorio de bases de datos debe contar con tecnología de virtualización, puesto que mejora los procesos. Desde el punto de vista de prácticas:

- el estudiante puede aprovechar el recurso de hardware virtualizado al 100%, mejorando así los tiempos de respuestas en una práctica.
- El estudiante podría desinstalar e instalar dentro de una máquina virtual todo el software requerido para desarrollar sus prácticas, sin tener ninguna restricción o por falta de permisos de usuario.
- El estudiante puede realizar copias de su máquina virtual y llevarlas consigo mismo y de esta manera trabajar desde cualquier parte.

³⁹ Eugenio Eduardo Villar Fernández. Virtualización de servidores de telefonía IP en GNU/Linux. Almería, 2010. Trabajo de grado (Ingeniero en Informática). Universidad de Almería.

Desde el punto de vista operativo:

- La universidad tendría ahorros en costos de energías por equipos adicionales como servidores.
- Ahorro en mantenimiento que se le debe dar a los equipos.
- Ahorro en equipos de refrigeración y entre otros.

Por estas razones, contar con herramientas de virtualización representa una mejor manera de proveer y gestionar servicios informáticos dentro de un laboratorio y más cuando se trata de un laboratorio de bases de datos enfocados para el aprendizaje de los estudiantes, representando un ahorro de dinero considerable en cuanto adquisición de equipos, espacio para ubicarlos, energías para ponerlos a trabajar, todo estos son aspectos importantes a considerar y tener en cuenta.

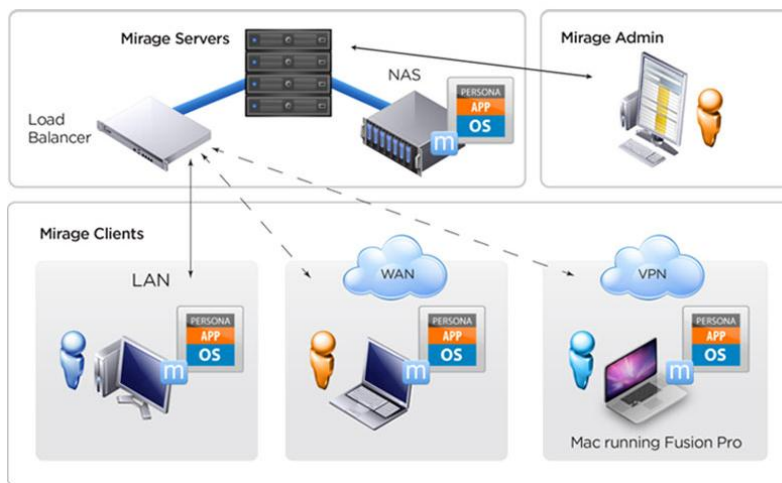
A continuación, nombraremos las herramientas más utilizadas para virtualización, seleccionando como vmware y virtualbox para el laboratorio de bases de datos, ya que estas ofrecen los requisitos para el propósito de las prácticas.

6.20.5.3 Herramientas de virtualización

- **Boch.** Es un emulador de arquitecturas basadas en X86 que funciona en múltiples plataformas, el mayor interés de Boch es que es capaz de emular un pc completo incluyendo los periféricos y funciona prácticamente cualquier sistema anfitrión.

- **Qemu.** Es un emulador que tiene dos modos de funcionamiento, uno de emulación de sistema completo y otra emulación en modo usuario. En el modo de sistema completo el programa emula el equipo entero incluyendo múltiples procesadores y periféricos, este modo se usa para ejecutar sistemas operativos completos, en las últimas versiones del programa se soportan más de 15 arquitecturas diferentes. En la emulación en modo usuario, el programa puede ejecutar programas compilados para una CPU concreta en un sistema que funciona sobre una CPU diferente.
- **VMware.** “VMware Horizon (con View) les permite a las organizaciones ofrecer escritorios virtualizados o remotos y aplicaciones a través de una única plataforma. Asimismo, les brinda respaldo a los usuarios finales para que puedan acceder a todos sus escritorios y aplicaciones en un solo lugar.”⁴⁰

Figura 5. Administración de imágenes con vmware



Fuente: VMware. [En línea] <<http://www.vmware.com/co/products/horizon-view/features.html>>

⁴⁰ VMware. [En línea] <<http://www.vmware.com/co/products/horizon-view/features.html>> [Citado en 21 de Octubre de 2014]

- **Microsoft VDI.**

Microsoft Virtualization Desktop Infrastructure (VDI), con tecnología de Windows Server 2012, le permite implementar arquitecturas de servicios de escritorio remotas que ofrecen a los empleados la flexibilidad de trabajar desde cualquier lugar, a la vez que les permite acceder de forma fluida a su entorno corporativo de aplicaciones o escritorio Windows ejecutado en el centro de datos desde una amplia gama de dispositivos. Las características y la infraestructura de administración unificada para escritorios centralizados en Windows Server 2012, combinadas con las tecnologías de virtualización de aplicaciones y estados de usuario con System Center, aumenta la flexibilidad del acceso para aplicaciones y escritorios remotos, ofreciendo unas experiencias personalizadas, coherentes y seguras a los usuarios, a la vez que mejoran el cumplimiento a través de un control y acceso centralizados a los datos confidenciales.⁴¹

- **Virtual box Oracle.**

Es un poderoso x86 y AMD64 / Intel64 virtualización de productos para la empresa, así como el uso doméstico. No sólo es VirtualBox extremadamente rico en características, producto de alto rendimiento para los clientes empresariales, es también la única solución profesional que está libremente disponible como software de código abierto bajo los términos de la Licencia Pública General de GNU (GPL) versión 2. Actualmente, VirtualBox se ejecuta en Windows, Linux, Macintosh, y los anfitriones de Solaris y soporta un gran número de sistemas operativos invitados incluyendo pero no limitado a Windows (NT 4.0, 2000, XP, Server 2003, Vista, Windows 7, Windows 8), DOS / Windows 3.x, Linux (2.4, 2.6 y 3.x), Solaris y OpenSolaris, OS / 2, y OpenBSD.⁴²

⁴¹ Microsoft. Virtualización del sistema operativo [En línea] <http://www.microsoft.com/es-es/windows/enterprise/products-and-technologies/virtualization/operating-system/default.aspx> [Citado en 21 de Octubre de 2014]

⁴² VirtualBox.[En línea] <<https://www.virtualbox.org/>> [Citado en 21 de Octubre de 2014]

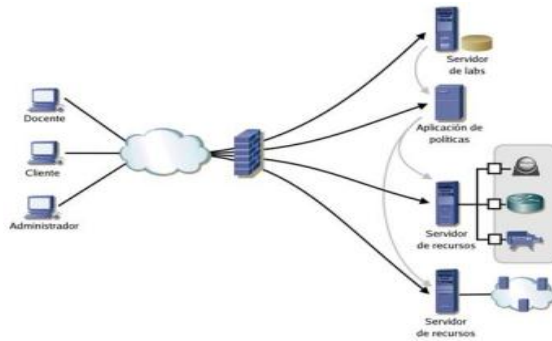
6.20.6 Acceso remoto al laboratorio de bases de datos

El acceso remoto al laboratorio de bases de datos, es de gran importancia para la comunidad estudiantil, ya que por medio de aplicaciones web la comunidad ingenieril pueden acceder a los recursos del laboratorio, permitiendo a los usuarios interactuar con las herramientas instalados en los equipos de la Universidad Católica de Colombia sin estar en contacto directo con ellos. Dada a la gran importancia que los laboratorios de acceso remoto ha tomado, se ha visto la necesidad de ser implementados en la educación debido a que son parte fundamental en la formación del estudiante, reforzando el conocimiento teórico aprendido por medio de la práctica.

El laboratorio para brindar servicio de acceso remoto contara con una infraestructura basada en modelos de laboratorios remotos, tales como (ver Figura 6):

- un servidor labs, que funciona como un portal para los usuarios, en este caso estudiantes, docentes y administradores. Su funcionalidad es que estos usuarios puedan iniciar una sesión segura.
- El servidor de recursos (SR), brinda acceso a recursos físicos y virtuales en una organización, este está conectado físicamente a los recursos.
- El servidor de políticas, es el encargado de permitir que cada usuario luego de ser autenticado tenga acceso a todos los recursos que brinda y ofrece el laboratorio con los permisos correspondiente.
- El servidor LDAP actúa como una base de datos que autentica cada usuario, para que pueda acceder al portal y por ende al laboratorio.

Figura 6. Esquema de acceso remoto



Fuente: E. Grosclaude, L. M. Bertogna, R. Del Castillo, F. López Luro y C. Zanellato, «sedici.unlp.edu.ar,» Available: http://sedici.unlp.edu.ar/bitstream/handle/10915/23272/Documento_completo.pdf?sequence=1.

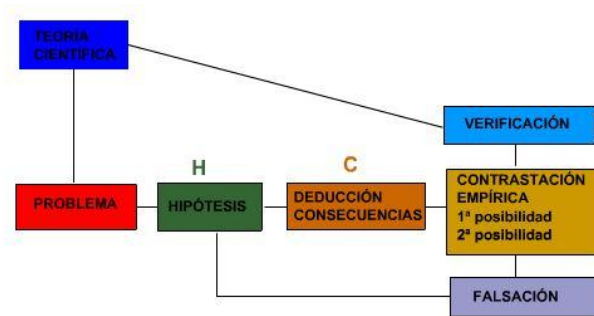
7. METODOLOGÍA

Este trabajo de grado sigue la metodología hipotético-deductiva. Desarrollada por varios pasos.

7.1 FASE DEL MÉTODO HIPOTÉTICO-DEDUCTIVO

En la Figura 7, se puede evidenciar los pasos del método hipotético-deductivo.

Figura 7. Método hipotético-deductivo



Fuente: I.E.S. Dionisio Aguado. Método Hipotetico-Deductivo. Disponible en: http://iesdionisioaguado.org/joomla/index.php?option=com_content&view=article&id=509:metodo-hipotetico-deductivo&catid=126:la-ciencia&Itemid=28

El método hipotético-deductivo posee tres fases para establecer una validez científica de la investigación que se explicaran a continuación:

7.1.1 Observación. Es la fase donde el investigador descubre el problema. Normalmente surge con dudas o inquietudes, sin embargo la observación también puede ser accidental. Por otro lado, para que adquiera el calificativo de científico tiene que reconocer un fenómeno que pueda ser medido o que se pueda contar (cuantificable).

7.1.2 Formulación de la hipótesis. Luego de ser realizada la observación el investigador procede a indagar una explicación. Para esto, el investigador elabora una teoría donde exponga todos los hechos encontrados. En cuanto los datos deducidos de la hipótesis tenga mayor veracidad en la experimentación y así la hipótesis será más certera. De este modo, una característica de las ciencias empíricas es que puede existir un caso particular donde se pueda llegar a refutar la validez de la hipótesis.

7.1.3 Deducción de las consecuencias de la hipótesis. Luego el investigador al construir la hipótesis debe derivar sus respectivas consecuencias que podría tener la hipótesis al momento de llegar a ser comprobada.

7.1.4 Contrastación o verificación de la hipótesis. “Una vez formulada la hipótesis y sus consecuencias es preciso proceder a su verificación o contrastación”⁴³.

⁴³ UNIVERSIDAD DE CANTABRIA. Etapas del método hipotético deductivo [en línea].
< <http://ocw.unican.es/ciencias-de-la-salud/ciencias-psicosociales-i/materiales/bloque-i/tema-1/1.1.3.1-etapas-del-metodo-hipotetico-deductivo> > [citado en 05 de agosto de 2014]

8. CRONOGRAMA

A continuación se presenta el cronograma de actividades que se desarrollara con este trabajo de grado.

Tabla 1. Cronograma de actividades

#	Actividades	Fecha inicio	Días de duración	Fecha de Finalización
1	Investigación del estado actual	20 de agosto de 2014	10	29 de agosto 2014
2	Levantamiento de requerimientos	29 de agosto de 2014	12	09 de septiembre de 2014
3	Diseño y desarrollo de la propuesta	09 de septiembre de 2014	29	07 de octubre 2014
4	Validación de resultados	7 de octubre 5 de octubre	21	21 de octubre 2014.
5	Entrega proyecto final	4 de noviembre de 2014		
6	Socialización proyecto	20 de noviembre de 2014		

Fuente: Los autores

9. ALCANCE Y LIMITACIONES

Este trabajo de grado tendrá los siguientes alcances y limitaciones:

9.1 ALCANCE

El alcance general de este proyecto está definido como el diseño de laboratorio de bases de datos para la universidad Católica de Colombia.

9.2 LIMITACIONES

Tiempo, acceso a la información, conocimientos del tema.

9.3 ESPACIO

El espacio donde se diseñara el proyecto de investigación se limita a la universidad católica de Colombia, donde se encuentra el sitio de interés para el desarrollo del laboratorio.

9.4 TIEMPO

El tiempo de realización de este proyecto con base a la planificación realizada por la Universidad Católica de Colombia inicia el 8 de Agosto con la formulación del anteproyecto y termina el 20 de noviembre con la sustentación de proyectos de grado.

10. ESTRATEGIAS DE COMUNICACIÓN

Este trabajo de grado tiene como estrategias de comunicación, en primera instancia, la socialización del trabajo de grado mediante una exposición clara sobre el proyecto, a su vez, se requiere elaborar un artículo científico que posteriormente podrá ser enviado a un congreso ya sea nacional o internacional. Por último, la memoria que se genera mediante este proyecto también es vista como estrategia de comunicación.

11.PRESUPUESTO

Tabla 2. Talento humano

Concepto	Tiempo (hr)	Costo(hr)	Subtotal
Documentación	50	\$ 20.000	\$ 1'000.000
Investigación	50	\$ 35.000	\$ 1'750.000
Diseño	80	\$ 40.000	\$ 3'200.000
Consultoría	20	\$ 50.000	\$ 1'000.000
Total	200	\$ 145.000	\$ 6'950.000

Fuente: Los autores

Tabla 3. Servicios

Concepto	Tiempo(mes)	Costo(mes)	Subtotal
Internet	4	\$ 98.000	\$ 392.000
Energía	4	\$ 50.000	\$ 200.000
Total	8	\$ 148.000	\$ 592.000

Fuente: Los autores

Tabla 4. Presupuesto global del proyecto

PRESUPUESTO GLOBAL DEL PROYECTO		
DESCRIPCIÓN	INGRESOS	EGRESOS
APORTES UNIVERSIDAD CATÓLICA DE COLOMBIA	\$ 4'200.000	
APORTES PROPIOS ESTUDIANTES	\$ 3'342.000	
Recurso Humano (32 Horas Docente)		\$ 4'200.000
Recurso Humano (432 Horas Estudiantes)		\$ 2'750.000
432 Horas de Equipo PC con Internet		\$ 592.000
Imprevistos (5%)		
TOTAL	\$ 7.542.000	\$ 7.542.000

Fuente: Los autores

12.LEVANTAMIENTO DEL ESTADO ACTUAL DEL LABORATORIO DE BASES DE DATOS

En este apartado del trabajo de grado, se va a elaborar el levantamiento de requerimientos para abordar el diseño del laboratorio de bases de datos para el Programa de Ingeniería de Sistemas de la Universidad Católica de Colombia. Se resalta que este proyecto está enmarcado en el plan de mejoramiento del programa para seguir fortaleciéndolo y por supuesto, para mejorar su currículo a través de herramientas prácticas, todo esto, con el fin de aportar hacia la comunidad académica (estudiantes y profesores) nuevos esquemas de aprendizaje para que desarrollen nuevas y mejores competencias para la profesión. El laboratorio de bases de datos para las carreras de Ingeniería de Sistemas o afines, es, en este momento, una herramienta académica fundamental, porque la información, el recurso más valioso de una organización requiere diversos esquemas de tratamiento y control. Además, el crecimiento de los sistemas de información y la gran cantidad de datos que se maneja actualmente generan la necesidad de que la Universidad y su programa, cuenten con esquemas de enseñanza para esta temática que, incluso, puede aprovecharse como herramienta que aportará hacia la generación de recursos para el programa y para la Universidad. Para poder abordar debidamente este planteamiento, es necesario analizar cuáles son los recursos de laboratorio que actualmente tiene el programa para que se parta de la base correcta y se genere un lineamiento claro que lleve a diseñar el nuevo laboratorio que se plantea mediante este trabajo de grado.

Para poder analizar correctamente toda la información asociada a este planteamiento, se precisa aclarar que mediante la Dirección de programa se pudo obtener información relevante de la forma en que está dispuesto su currículo y obviamente los laboratorios asociados al programa.

Actualmente el programa de ingeniería de sistemas, tiene definidas siete (7) áreas del conocimiento que comprenden 27 asignaturas y cada una desarrolla competencias específicas de la profesión.

Tabla 5. Áreas del conocimiento

Área conocimiento	No. Asignaturas
Modelación matemática	3
Ciencias de la computación	6
Programación y estructura de datos	5
Administración de la información	4
Ingeniería de software	6
Sistemas y organizaciones	5
Redes	4

Fuente: Los autores

Este proyecto de grado, se centra específicamente en el área de la administración de la información, compuesta por 4 asignaturas como se puede observar en la Tabla 6.

Tabla 6. Asignaturas área administración de la información

ARQUITECTURADELCOMPUTADOR	CT10055	Administración de la información
FUNDAMENTOS DE BASES DE DATOS	CT10002	
SISTEMAS OPERATIVOS	CT10022	
INTELIGENCIA DE NEGOCIOS *	CT10081	

*Electiva del programa

Fuente: Suministrado por el programa de Ingeniería de sistemas

Por su parte, es de resaltar que este laboratorio no únicamente apoya y aporta a estas asignaturas sino también que transversalmente apoya al área de ingeniería de software, en especial, a las asignaturas de sistemas de información, ingeniería de software y arquitectura de software.

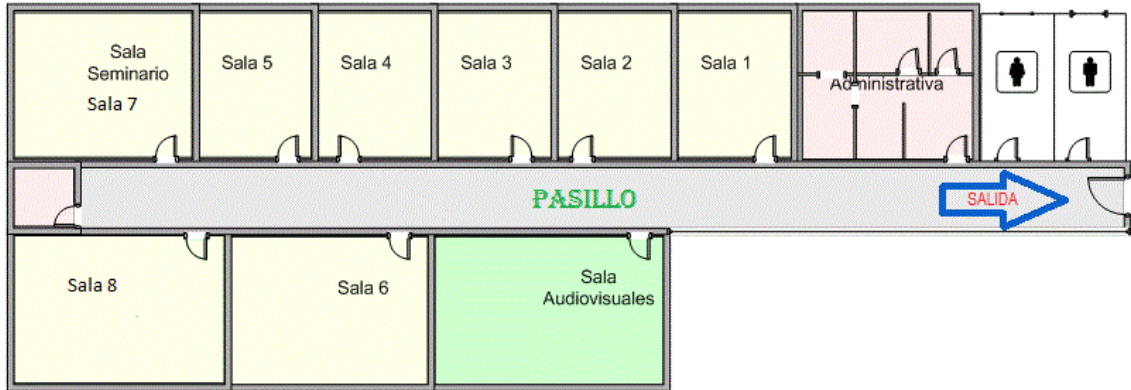
Tomando de referencia los documentos suministrados por el programa para el desarrollo de este trabajo de grado, es importante mencionar que, en su estructura curricular, se menciona que las asignaturas asociadas al área de administración de la información están enfocadas a generar “conocimientos en la gestión de la información en las organizaciones, y dan lineamientos claros del qué, por qué, para qué, cuando y donde deben utilizarse. A su vez, es importante resaltar que en las asignaturas no solamente se contempla la información electrónica sino la física porque todo hace parte activa de la organización y además, el ingeniero debe entender como es la estructura organizacional para que sea capaz de manejar los datos referentes a las distintas instancias que componen la organización sin importar la fuente o el formato (datos, documentos en papel, documentos electrónicos, audio, video, imágenes, etc.), todo esto, con el fin de poder entregar la información a través de múltiples canales que puedan incluir diversas interfaces tecnológicas para apoyar el manejo de información de la empresa.

Por otra parte, para lograr el aprendizaje óptimo de las asignaturas que hacen parte del área en mención, la Universidad Católica de Colombia posee un laboratorio de sistemas, ubicado en el tercer piso del bloque O de la sede claustro, distribuido en ocho (8) salas (ver Figura 8) y una sala denominada Samsung training Center ubicada en el primer piso. Tales salas cuentan con una serie de equipos de cómputo con sus respectivas características (ver Tabla 8) y una capacidad promedio de 20 a 25 estudiantes por sala (ver Tabla 7). Es de resaltar que esta información fue obtenida a través de la coordinación de laboratorio de sistemas.

A continuación se presenta la información referente a los laboratorios, en especial, a las salas de cómputo que utiliza el programa de ingeniería de sistemas y otros programas.

Figura 8. Mapa laboratorio de sistemas

MAPA LABORATORIO DE SISTEMAS



Fuente: PARDO P. Bernardo. Informe estado actual de las salas para el programa de ingeniería de sistemas.

De acuerdo con el estado actual de las salas, a continuación se presenta la disponibilidad y capacidad de las salas que utiliza el programa de ingeniería de sistemas y otros programas.

Tabla 7. Capacidad y disponibilidad salas

Laboratorio	Sala 1	Sala 2	Sala 3	Sala 4	Sala 5	Sala 6	Sala 7	Sala 8		Audiovisuales
	Modelos	DELL 780	DELL 780	DELL 780	DELL 780	DELL 780	DELL 9010	DELL 780	DELL 780	DELL 9010
Disponibilidad de Equipos	16	16	15	17	15	30	15	4	16	1
Capacidad Estudiantes / Sala	20	20	20	20	25	35	30	30		36

Fuente: PARDO P. Bernardo. Informe estado actual de las salas para el programa de ingeniería de sistemas.

Posteriormente, se presenta las características de hardware de los equipos que utiliza el programa de ingeniera de sistemas y otros programas.

12.1 HARDWARE Y SOFTWARE ACTUAL

12.1.1 Características de equipos de cómputo

Tabla 8. Características equipos de cómputo

		DELL		
		Optiplex 745/755	Optiplex 780	Optiplex 9010
Distribucion de Equipos de Computo	Carácter.			
	Disco Duro:	80 Gb	160 Gb	320 gb
	Memoria:	2 Gb	4 Gb	8 gb
	Procesador:	Intel Core 2 Duo 2.3 GHZ	Intel Core 2 Duo 3 GHZ	Intel cor i5
S.O.:	Fedora 18/Linux Centos	Win 7 PRO. De 32 bit.	Win 7 PRO. De 64 bit	

Fuente: PARDO P. Bernardo. Informe estado actual de las salas para el programa de ingeniería de sistemas.

A continuación, se describen las características que poseen de los servidores que se encuentran ubicados en la sala de administración del laboratorio de ingeniera de sistemas de la Universidad Católica de Colombia.

12.2 CARACTERÍSTICAS DE SERVIDORES

La sala de administración de laboratorio cuenta con una serie de servidores utilizados para propósito de las prácticas y se encuentran distribuidos de la siguiente forma: 9 servidores, de los cuales 4 están dedicados al proyecto GridComputing, un servidor de dominio para todas las salas (ver Tabla 9), el cual se encarga de administrar los usuarios, crear políticas de grupos y la no instalación de software ni el ingreso de un estudiante con un usuario diferente al de la sala en la cual se encuentra; todas estas políticas son establecidas por el área de informática que está ubicada en la sede de las torres.

A continuación se presenta la descripción de cada servidor nombrados anteriormente.

Tabla 9. Características servidor controlador de dominio

Servidor Controlador de Dominio		
Servidor Dell Poweredge 2950		
Xeon 2.4 GHz	Procesador	1
3 Discos de 146 GB	Disco duro	2
4 GB	Memoria Ram	1
Windows 2008 Server	Sistema Operativo	1

Fuente: PARDO P. Bernardo. Informe estado actual de las salas para el programa de ingeniería de sistemas.

Un servidor de espejo de dominio (ver Tabla 10) para crear redundancia, en dado caso que haya una tipo de falla en el servidor principal.

Tabla 10. Características servidor espejo

Servidor Controlador de Dominio Espejo		
Servidor Dell PowerEdge R320		
Intel® Xeon® E5-2440 Dual core 2.4 GHz	Procesador	1
2 Discos de 500 GB	Disco duro	2
8 GB	Memoria Ram	1
Windows 2012 Server	Sistema Operativo	1

Fuente: PARDO P. Bernardo. Informe estado actual de las salas para el programa de ingeniería de sistemas.

Un servidor web (ver Tabla 11) con dirección IP pública para poder acceder desde internet.

Tabla 11. Características servidor web apache

Servidor WEB APACHE-TOMCAT		
Servidor Dell PowerEdge R320		
Intel® Xeon® processor E5-2400 product	Procesador	1
2 Discos de 500 GB	Disco duro	2
8 GB	Memoria Ram	1
Centos ver 6.4	Sistema Operativo	1

Fuente: PARDO P. Bernardo. Informe estado actual de las salas para el programa de ingeniería de sistemas.

También cuenta con un servidor Centos (ver Tabla 12) el cual virtualiza servicios de Oracle con licencia de renovación anual con fines académicos.

Tabla 12. Características servidor clúster

	Carácter.	Capacidad
Característica de Servidores (Cluster)	Disco Duro:	500 Gb
	Memoria:	8 Gb
	Procesador:	Xeon 2.4 GHZ
	S.O.:	-----
	Modelo:	DELLR310
	Marca:	DELL
	Cantidad:	4
	S.O.:	CentOs 6,4

Fuente: PARDO P. Bernardo. Informe estado actual de las salas para el programa de ingeniería de sistemas.

Un servidor Sun Fire v.245 (Ver Tabla 13) que tiene instalado Oracle 10g, para que los estudiantes realicen las prácticas en bases de datos.

Tabla 13. Características servidor base de datos

Servidor Bases de Datos Oracle		
Servidor Sun Fire V245		
Ultra Sparc IIIC 1.5 GHz	Procesador	1
2 Discos de 73 GB	Disco duro	2
4 GB	Memoria Ram	1
Solaris 10.0	Sistema Operativo	1

Fuente: PARDO P. Bernardo. Informe estado actual de las salas para el programa de ingeniería de sistemas.

La distribución de los servidores en mención, se realiza dentro de un rack (ver Figura 9) localizado en la sala de administración de laboratorio. Se resalta que el servicio de internet es proporcionado por la sala de informática, ubicada en la sede de las torres, en el primer piso, donde está el centro de procesamiento de datos, que brinda gestión y administración al laboratorio ubicado en la sede de la carrera

13 y también apoyo a otros laboratorios como electrónica, arquitectura, psicología y ciencias básicas.

Figura 9. Servidores U. Católica



Fuente: Los autores

Este centro de procesamiento de datos, cuenta con piso falso (ver Figura 10), control de acceso para el ingreso del personal, cuarto eléctrico donde se encuentran alojadas las ups, detector de incendios (ver Figura 11) y un sistema de iluminación y refrigeración. A su vez cuenta con un rack de servidores (ver Figura 12) que está destinado a todos los trabajos críticos que se realizan en la universidad.

Figura 10. Piso falso - sala informática



Fuente: Los autores

Figura 11. Sistema de detección de incendios



Fuente: Los autores

Figura 12. Servidores sala informática de la Universidad Católica de Colombia



Fuente: Los autores

12.3 SOFTWARE ACTUAL

Cada una de las salas del laboratorio del programa de ingeniería de sistemas cuentan con diverso software (ver Figura 13), sin embargo, dependiendo de la práctica a realizar se elige la sala donde es posible realizar tal práctica y posteriormente se le informa al docente y a vez al estudiante. Es de resaltar que ninguna de estas salas están preparadas para las áreas curriculares como tal, pero responden medianamente a las necesidades del programa.

Figura 13. Software disponible en salas

Software	Version	Licencia	Cant. C/S	Salas de Computo									
				Admon	Sala 1	Sala 2	Sala 3	Sala 4	Sala 5	Sala 6	Sala 7	Sala 8	
S.O.	Windows 7	Profesional	Convenio *	153/0	X	X	X	X	X	X	X	X	X
	Windows 8	Pro-64	Convenio *	17/0	X	---	X	---	---	---	---	---	---
	Windows XP	Profesional SP2	Convenio *	154/0	X	X	X	X	X	X	X	X	X
	Windows 2003	Server	Convenio *	0/1	X	---	---	---	---	---	---	---	---
	Windows 2012	Server	Convenio *	0/3	X	---	---	---	---	---	---	---	---
	Linux Ubuntu	12.04 LTS	Libre	1	X	---	---	---	---	---	---	---	---
	Unix Solaris	10	Licenciado	0/1	X	---	---	---	---	---	---	---	---
	Linux Centos	5	Libre	0/1	X	---	---	---	---	---	---	---	---
	Linux Centos	6.4	Libre	1/0	X	---	---	---	---	---	---	---	---
	Linux Scientific	6	Libre	30/0	---	---	---	X	---	X	---	---	---
	Linux Fedora	18	Libre	53/0	---	---	---	---	X	---	---	X	X
Adobe	Adobe Flash Plug-in	11.8.800.94	Plug-in Libre	154	---	X	X	X	X	X	X	X	X
	Adobe Shockwave	11.6.8.638	Plug-in Libre	154	---	X	X	X	X	X	X	X	X
	Adobe Reader	11.0.5	Plug-in Libre	154	---	X	X	X	X	X	X	X	X
	Adobe Air	3.5.0.1060	Plug-in Libre	154	---	X	X	X	X	X	X	X	X
Microsoft	Microsoft Office	2013	Convenio *	154	---	X	X	X	X	X	X	X	X
	Microsoft Silverlight	5.1.20513.0	Libre	154	---	X	X	X	X	X	X	X	X
	Visual Studio Premium	2012	Convenio *	154	---	X	X	X	X	X	X	X	X
	Complementos	---	Libre	154	---	X	X	X	X	X	X	X	X
Programacion - Lenguaje	Dev C++	4.9.9.2	Libre	154	---	X	X	X	X	X	X	X	X
	Eclipse	Kepler	Libre	154	---	X	X	X	X	X	X	X	X
	Java	7 update 25	Libre	154	---	X	X	X	X	X	X	X	X
	Jcreator LE	4.50	Libre	154	---	X	X	X	X	X	---	X	X
	Astah Community	6.7	Libre	154	---	X	X	X	X	X	X	X	X
	Net Beans	7.2	Libre	154	---	X	X	X	X	X	X	X	X
	Pseint	20130930-W32	Libre	154	---	X	X	X	X	X	X	X	X
	Zinjal	W32-20130801	Libre	154	---	X	X	X	X	X	X	X	X
Sistemas	SQL Plus (Oracle)	10g	Licenciado	500/1	X	---	X	X	X	---	---	---	---
	Oracle Express	10g	Libre	16	X	---	X	---	---	---	---	---	X
	MySQL	server 5.5	Libre	154	X	---	X	X	X	X	---	X	X
	Apache	2.2.16	Libre	154	X	X	X	X	X	X	X	X	X
	PHP	5.2.6	Libre	154	X	X	X	X	X	X	X	X	X
	Tomcat	7.0.27	Libre	154	X	X	X	X	X	X	X	X	X
Diagramas & Diseño	Cmap tools	5.04	Libre	154	X	X	X	X	X	X	X	X	X
	BizAgi	2.4.0.8	Demo	112	---	X	X	X	X	X	---	X	X
	Derive	5	Licenciado **	40/1	---	---	---	X	---	---	---	---	---
	Auto CAD	2014	Licenciado	30	---	---	---	---	---	---	X	---	---
	Cisco Packet Tracer	6.0.1	Libre	148	---	X	X	X	X	X	X	X	X
	Auto Desk	2014-4.0.32.100	Licenciado	30	---	---	---	---	---	---	X	---	---
	Corel Draw Studio	11	Licenciado	30	---	---	---	---	---	---	X	---	---
Otros	IZarc	3.6	Libre	154	---	X	X	X	X	X	X	X	X
	Virtual Box (Oracle)	4.2.18	Libre ***	154	X	X	X	X	X	X	---	X	X
	Google Chrome	30.0.1599.69	Libre	154	---	---	---	---	---	---	---	---	---
	MatLab	7.0	Licenciado **	40/1	X	---	---	---	X	---	---	X	---
Anti-Virus	Microsoft Forefront Endpoint Protection	4.2.223.0	Libre	154	X	X	X	X	X	X	X	X	X
	Ccleaner	4.06.4324	Libre	154	X	X	X	X	X	X	X	X	---
	Microsoft Essentials	4.3.216.0	Libre	5	X	---	---	---	---	---	---	---	---
Observaciones													
				* Estas licencias estan establecidas por convenio del Campus Agreement.									
				** El(los) Software(s) se encuentra activo pero desactualizado.									
				*** Licenciado, pero el software se puede utilizar de forma educativa en Colegios y Universidades									

Fuente: PARDO P. Bernardo. Informe estado actual de las salas para el programa de ingeniería de sistemas.

12.4 USO DE LAS SALAS

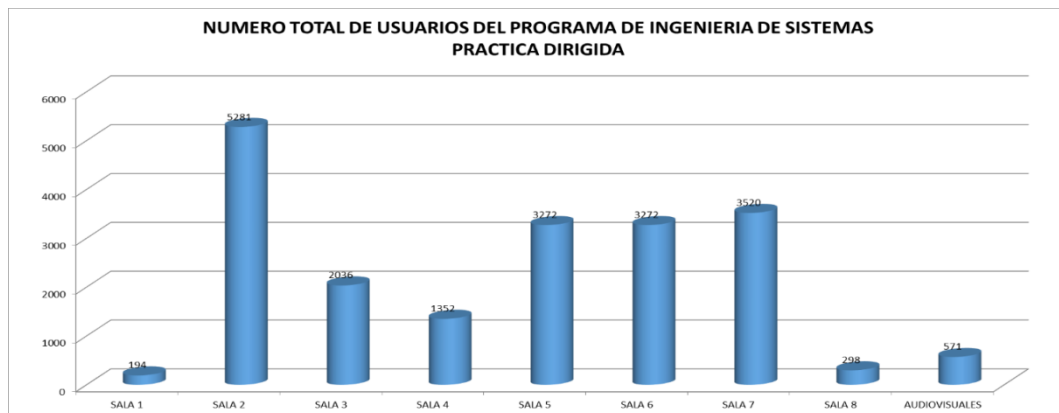
El horario del laboratorio del programa de ingeniería de sistemas es de lunes a viernes de 6:00 am a 10:00 pm y los sábados de 6:00 am a 6:00 pm, en jornada continua. Es de gran importancia conocer el uso que se le han dado a cada una de las salas, tanto para prácticas dirigidas como para las prácticas libres, por tal

motivo, a continuación se puede observar una serie de estadísticas del laboratorio del programa de ingeniería de sistemas en el primer semestre del año en curso (2014-I).

La utilización de las salas está dada por la realización de dos tipos de práctica: dirigida y libre. La primera es orientada por el docente como apoyo a las temáticas a desarrollar en la asignatura y la segunda es programada por el estudiante para ampliar sus conocimientos y desarrollar sus tareas asignadas.

12.4.1 Prácticas dirigidas. A continuación, se observa las estadísticas de la cantidad de estudiantes que hacen prácticas libres en las diferentes salas del laboratorio del programa de ingeniería de sistemas, en el primer semestre del año en curso (2014-I). El análisis que se puede desarrollar del gráfico (ver Figura 14), se puede observar que las salas con mayor uso para las prácticas dirigidas son; la sala 2, 5, 6 y sala 7, siendo la sala 2 la de mayor uso para el desarrollo de prácticas dirigidas, mientras que las salas 1 y 8 son las que menos son utilizadas para las prácticas dirigidas.

Figura 14. Número total usuarios prácticas dirigidas

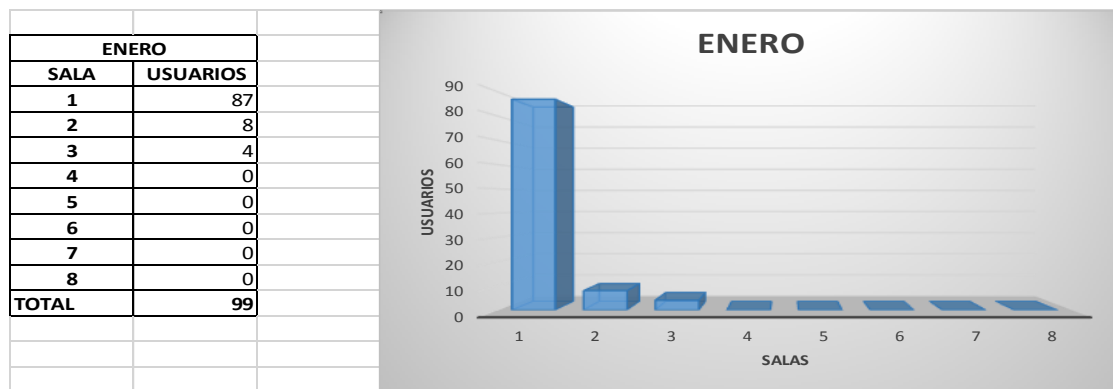


Fuente: PARDO P. Bernardo. Informe estado actual de las salas para el programa de ingeniería de sistemas.

12.4.2 Prácticas libres. A continuación, se podrá observar las estadísticas de la cantidad de estudiantes que hacen prácticas libres en las diferentes salas del laboratorio del programa de ingeniería de sistemas, en el primer semestre (ver Figura 15, Figura 16, Figura 17, Figura 18 y Figura 19) del año en curso (2014-I) por meses.

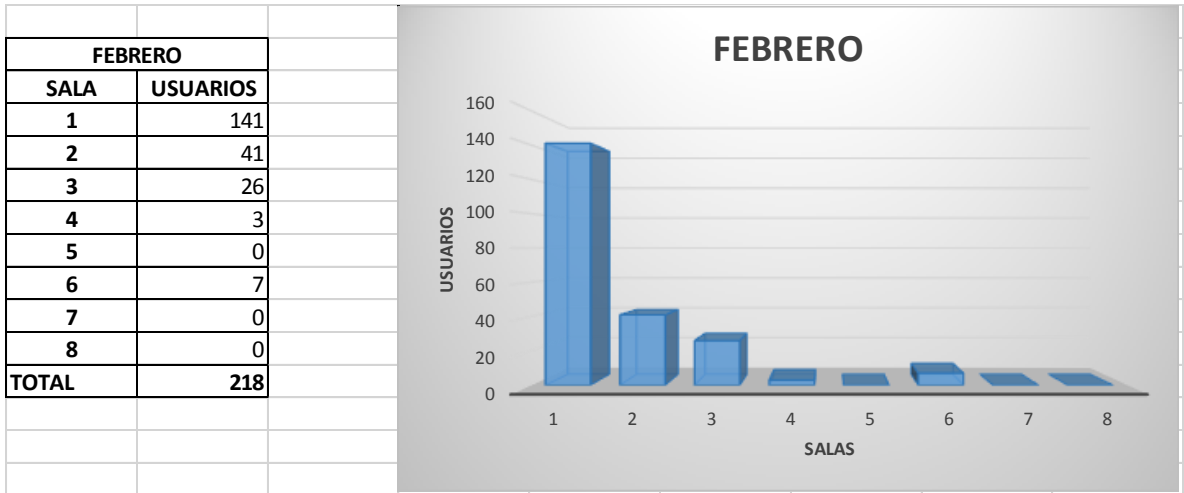
Haciendo un análisis del uso de cada una de las salas para las prácticas libres, es evidente notar que la sala con mayor uso es la sala 1, mientras que el uso de las demás salas para las prácticas libres se puede decir que es casi nulo.

Figura 15. Prácticas libres – Enero



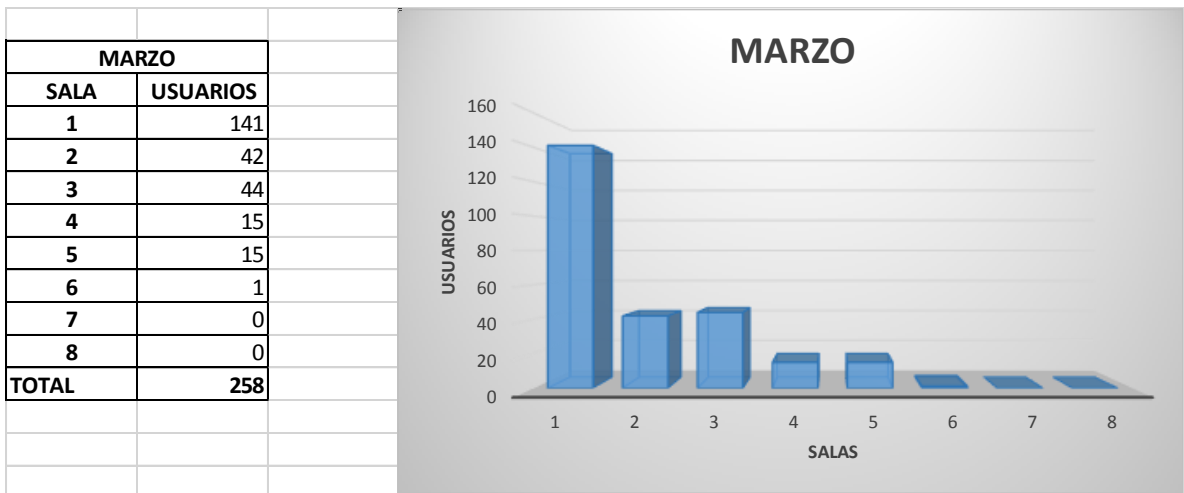
Fuente: PARDO P. Bernardo. Informe estado actual de las salas para el programa de ingeniería de sistemas.

Figura 16. Prácticas libres - Febrero



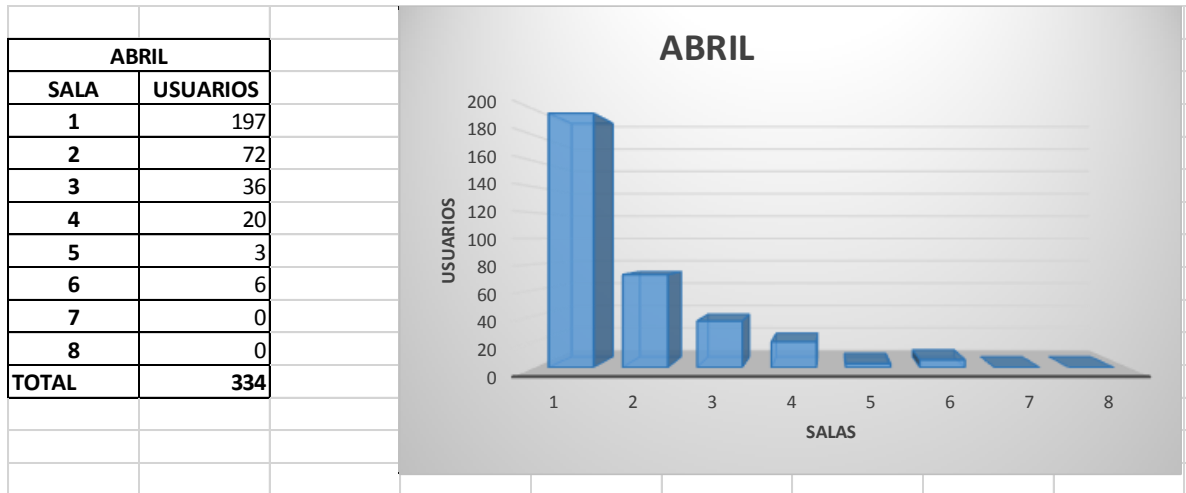
Fuente: PARDO P. Bernardo. Informe estado actual de las salas para el programa de ingeniería de sistemas

Figura 17. Prácticas libres – Marzo



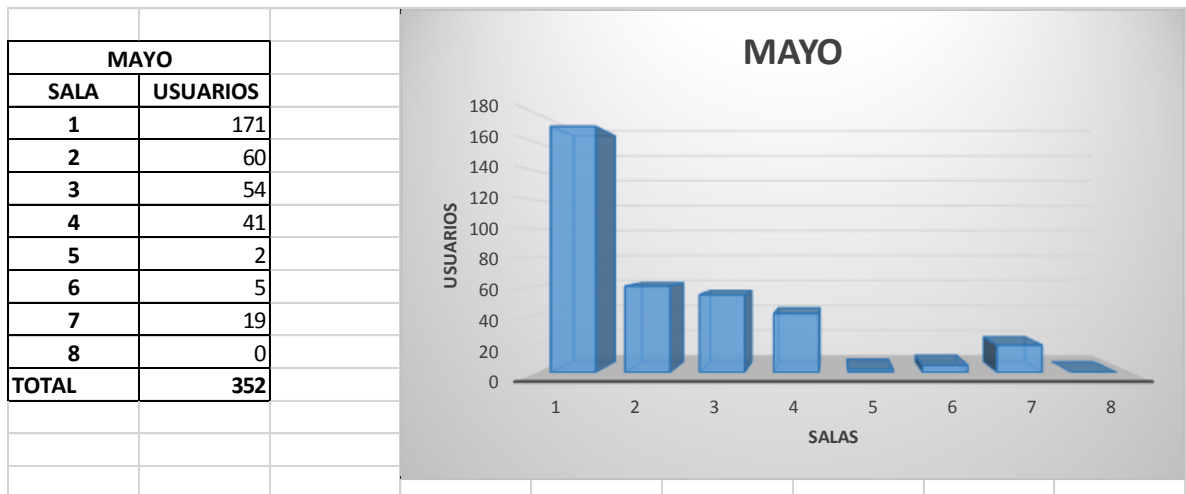
Fuente: PARDO P. Bernardo. Informe estado actual de las salas para el programa de ingeniería de sistemas.

Figura 18. Prácticas libres - Abril



Fuente: PARDO P. Bernardo. Informe estado actual de las salas para el programa de ingeniería de sistemas

Figura 19. Prácticas libres – Mayo



Fuente: PARDO P. Bernardo. Informe estado actual de las salas para el programa de ingeniería de sistemas

Haciendo una comparación en el uso de las salas para las prácticas dirigidas y las prácticas libres, es de notar que para las prácticas dirigidas hay un mayor uso de todas las salas mientras que para las prácticas libres la única sala asignada a los estudiantes es la sala 1, siendo esta la de mayor concurrencia.

Se precisa mencionar que de acuerdo a la información suministrada por la coordinación de laboratorio, cada sala en mención, es asignada a los docentes de acuerdo a los requerimientos de la asignatura. Analizando la documentación proporcionada, en el período 2014-I, las salas más utilizadas y orientadas a satisfacer el área curricular de la administración de la información fueron las salas 2 y 4 (ver Tabla 14), las cuales tuvieron un total de 104 sesiones durante el transcurso de este periodo, con un total aproximado de 212 horas de uso en solo el área de la administración de la información, dado que la sala 2 es la óptima en cumplir con los requerimientos de software y la única que cuenta con la herramienta Oracle Express, para el desarrollo de las prácticas de bases de datos.

Tabla 14. Análisis de uso de las sala 2 y la sala 4

Asignatura	Programa	Docente	Horas de uso	Sala utilizada	Sesiones	Jornada
Fundamentos de bases de datos	Sistemas	Carlos Pulido	54	2	27	Diurna
Inteligencia de negocios	Sistemas	Carlos Pulido	60	4	30	Diurna
Inteligencia de negocios	Sistemas	Carlos Pulido	66	4	31	Nocturno
Sistemas operativos	Sistemas	Miguel Dario Dussan	32	2	16	Nocturno
Sistemas operativos	Sistemas	Juan Carlos Navarro	2	Audio Visuales	1	Diurna

Fuente: Los autores

13.LEVANTAMIENTO DE REQUERIMIENTOS DEL LABORATORIO DE BASE DE DATOS

Este apartado tiene como fin desarrollar el levantamiento de requerimientos para el diseño del laboratorio de bases de datos, el cual se planteará a partir de las necesidades observadas. Para la recolección de datos se tomó una metodología de diseño, que contempla la opinión del docente del programa ya que ellos tienen conocimientos en el área de la administración de la información, razón por la cual se elaboraron dos encuestas. La primera (ver **Anexo A**) se aplicó a la población de docentes que dictan asignaturas del programa y la segunda **Anexo B**, fue dirigida específicamente a los que dictan asignaturas a fines a este laboratorio de bases de datos.

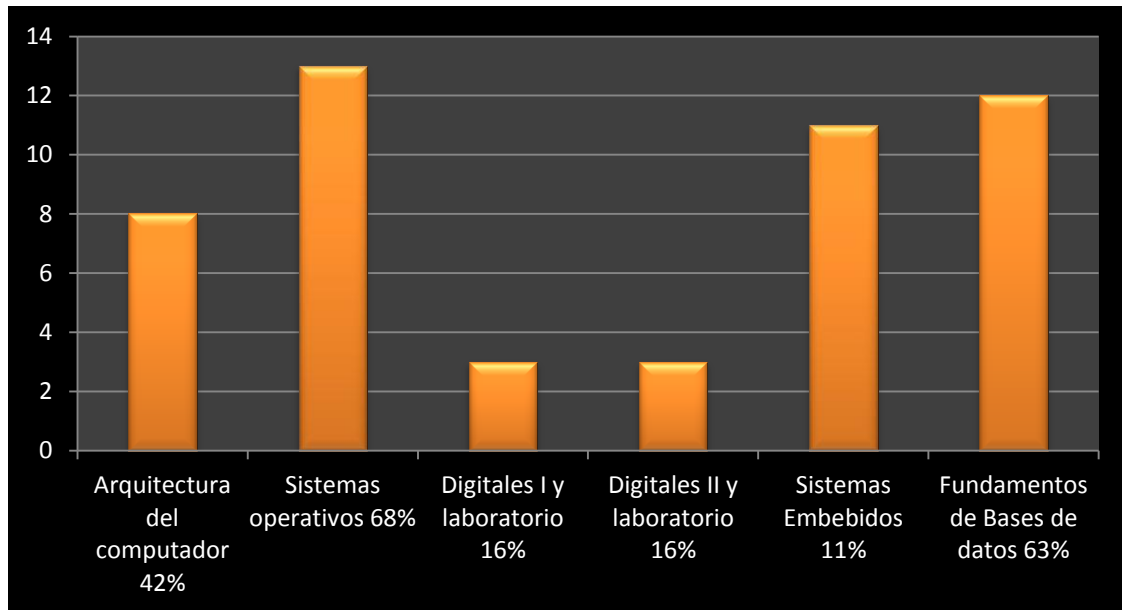
Estas encuestas consistieron en la elaboración y aplicación de 17 preguntas en total de selección múltiple a los docentes de la universidad católica de Colombia, los cuales accedieron a ella a través de un link enviado por correo electrónico; el análisis de estas respuestas se realizó a través de la interpretación de los resultados obtenidos con el fin de identificar los requerimientos necesarios para el diseño del laboratorio como apoyo al plan de mejoramiento en el que se encuentra la universidad.

A continuación, se presenta las encuestas realizadas a los docentes.

13.1 ENCUESTA GENERAL REALIZADAS A LOS DOCENTES

13.1.1 ¿Cuáles materias, según su enfoque académico, requieren de un incremento en la infraestructura de TI, para completar el desarrollo de las prácticas actuales?

Figura 20. Pregunta #1 desarrollada en la encuesta general



Fuente: Los autores.

- **INTERPRETACIÓN**

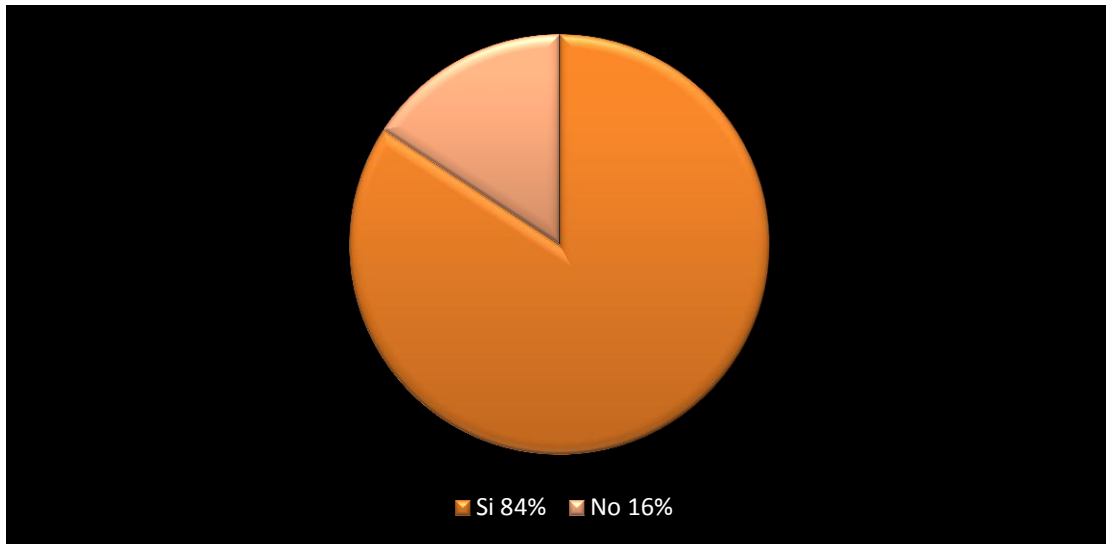
Los datos de la Figura 20, nos muestra un total de 6 materias que conforman el área de administración de la información, donde el 68% corresponde a 13 docentes que contestaron, los cuales tienen competencias en este tema, y solo dos de ellos dictan la asignatura de sistemas operativos, la cual necesita mayor incremento en la infraestructura TI y un menor porcentaje del 16% que corresponde a 3 docentes en digitales I y 3 docentes en digitales II para complementar el desarrollo de las prácticas actuales.

- **CONCLUSIÓN PARCIAL**

Se debería mejorar la infraestructura tanto de hardware como de software y entre otras que apoyen el despliegue a fondo de las prácticas actuales.

13.1.2 ¿Cree usted que los laboratorios requiere de accesos remotos para la realización de las prácticas?

Figura 21. Pregunta #2 desarrollada en la encuesta general



Fuente: Los autores

- **INTERPRETACIÓN**

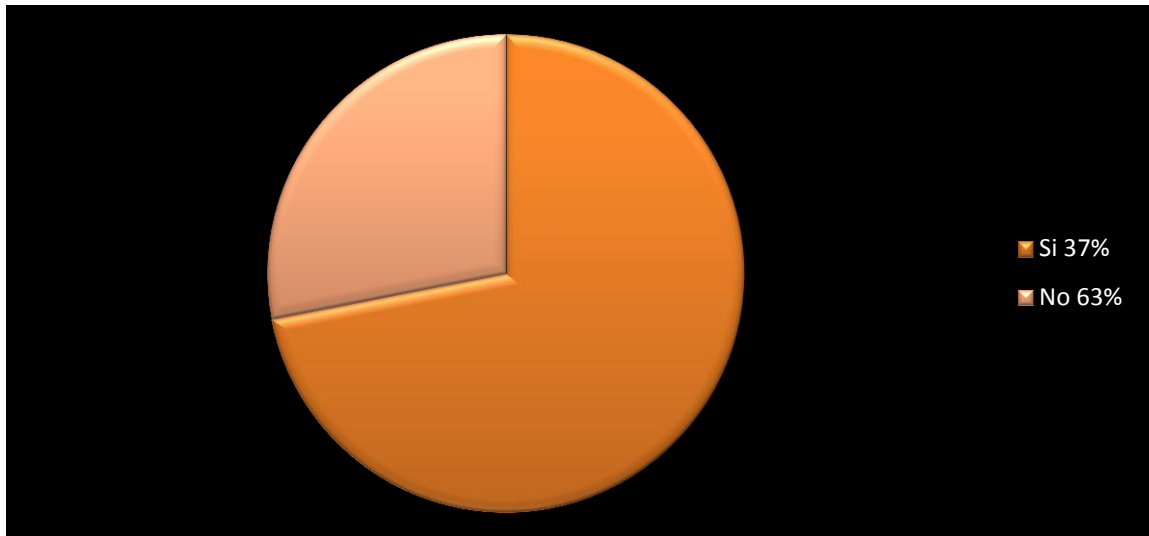
Según la Figura 21, requerir acceso remoto, para la realización de las prácticas, entre los docentes encuestados es de 84% que corresponde a 16 personas que ven con prioridad este ítem y un 16% que corresponde a 3 personas que no ven con importancia este componente dentro de los laboratorios.

- **CONCLUSIÓN PARCIAL**

Es de suma importancia que el laboratorio de bases de datos cuente con acceso remoto para la realización de las prácticas y de esta forma el estudiante aprovechar su tiempo libre y pueda acceder desde cualquier lugar.

13.1.3 ¿En los laboratorios se cuenta con la posibilidad de instalar software que permitan el desarrollo de las prácticas en diferentes ambientes?

Figura 22. Pregunta #3 desarrollada en la encuesta general



Fuente: Los autores.

- **INTERPRETACIÓN**

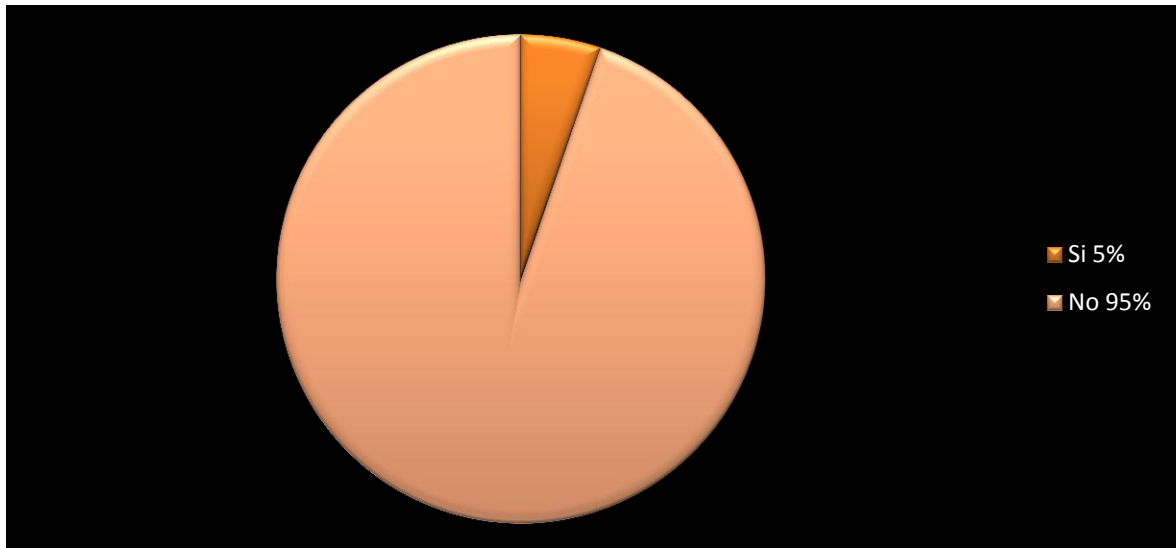
Segun la Figura 22, contar con la posibilidad de instalar software que permitan el desarrollo de prácticas en diferentes ambientes, entre los docentes encuestados es de un 63% equivalente a 12 personas que no ven la necesidad y con un 37% que corresponde a 7 personas, ven la importancia de poder hacer .

- **CONCLUSIÓN PARCIAL**

Es necesario que los estudiantes y profesores puedan instalar software requerido para poder desplegar las prácticas completamente.

13.1.4 ¿Los laboratorios cuentan con componentes de acceso remoto que amplien la disponibilidad, y apoyen el desarrollo de prácticas tanto dirigidas como libres?

Figura 23. Pregunta #4 desarrollada en la encuesta general



Fuente: Los autores

- **INTERPRETACIÓN**

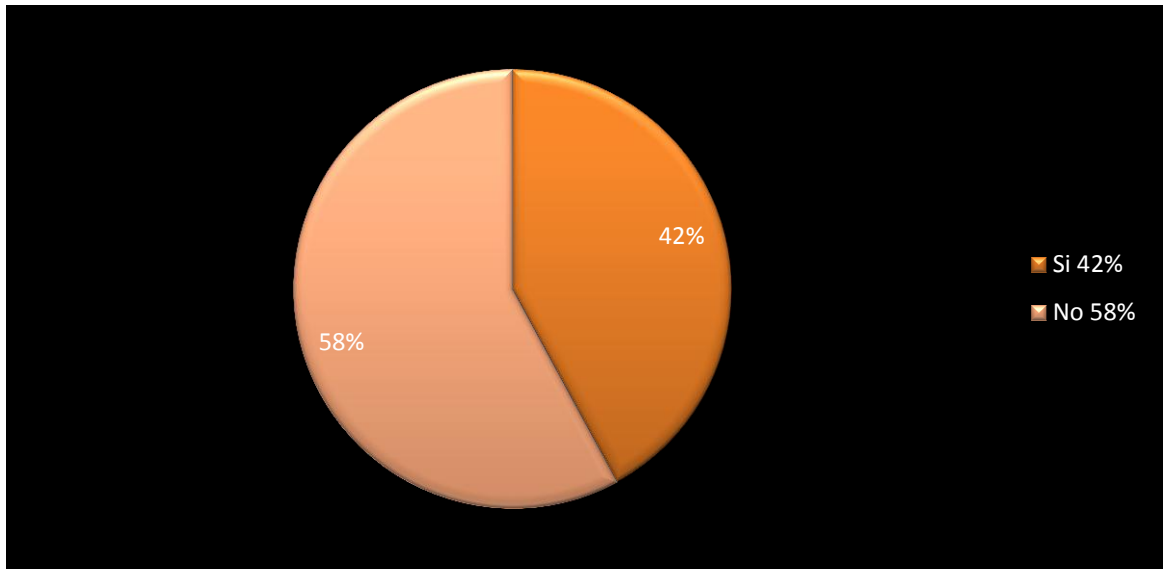
Según la Figura 23, los encuestados ven con suma importancia contar con componentes de acceso remotos que amplien la disponibilidad y apoyen el desarrollo de prácticas libres como dirigidas, .

- **CONCLUSIÓN PARCIAL**

Es de gran relevancia que el laboratorio de bases de datos cuente con acceso remoto, con el fin de apoyar al estudiante tanto en prácticas dirigidas como libres, brindandoles espacios extras para resolver dudas y desarrollar sus labores estudiantiles.

13.1.5 ¿Cuenta con los permisos de usuario necesarios para la realización de las prácticas en su totalidad?

Figura 24. Pregunta #5 desarrollada en la encuesta general



Fuente: Los autores

• INTERPRETACIÓN

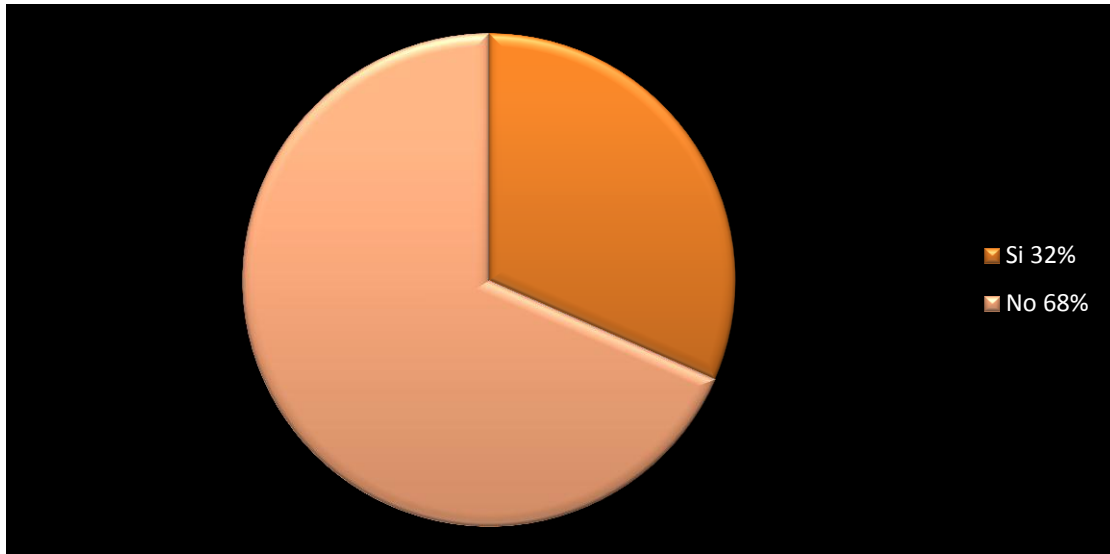
Segun la Figura 24, los encuestados ven con suma importancia, contar con los permisos de usuarios necesarios para la realizacion de las prácticas en su totalidad.

• CONCLUSIÓN PARCIAL

Es de gran importancia que los usuarios tanto estudiantes como profesores cuenten con permisos necesarios de usuarios para poder llevar acabo sus prácticas en su totalidad.

13.1.6 ¿Los laboratorios cuentan con el componente de virtualización?

Figura 25. Pregunta #6 desarrollada en la encuesta general



Fuente: Los autores

- **INTERPRETACIÓN**

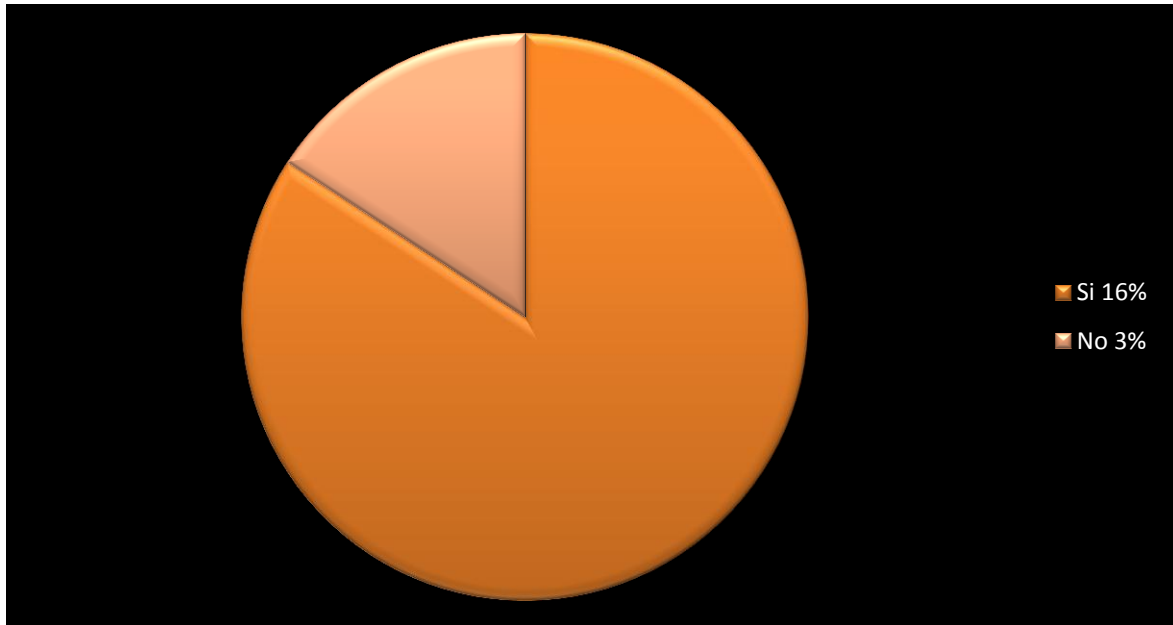
Según la Figura 25, contar con componentes de virtualización en los laboratorios, entre los docentes encuestados es de 32% que corresponde a 6 personas que respondieron que la universidad si cuenta con este componente y un 68% que corresponden a 13 personas que no lo encuentro.

- **CONCLUSIÓN PARCIAL**

Es necesario que los laboratorios cuenten con componentes de virtualización con el fin de aprovechar los recursos de hardware y de esta manera desarrollar mejores prácticas.

13.1.7 Teniendo presente la disposición actual de las salas, bajo su criterio ¿cree que se necesita un mayor espacio para el número promedio de estudiantes que acceden a este para realizar las prácticas?

Figura 26. Pregunta #7 desarrollada en la encuesta general



Fuente: Los autores.

- **INTERPRETACIÓN**

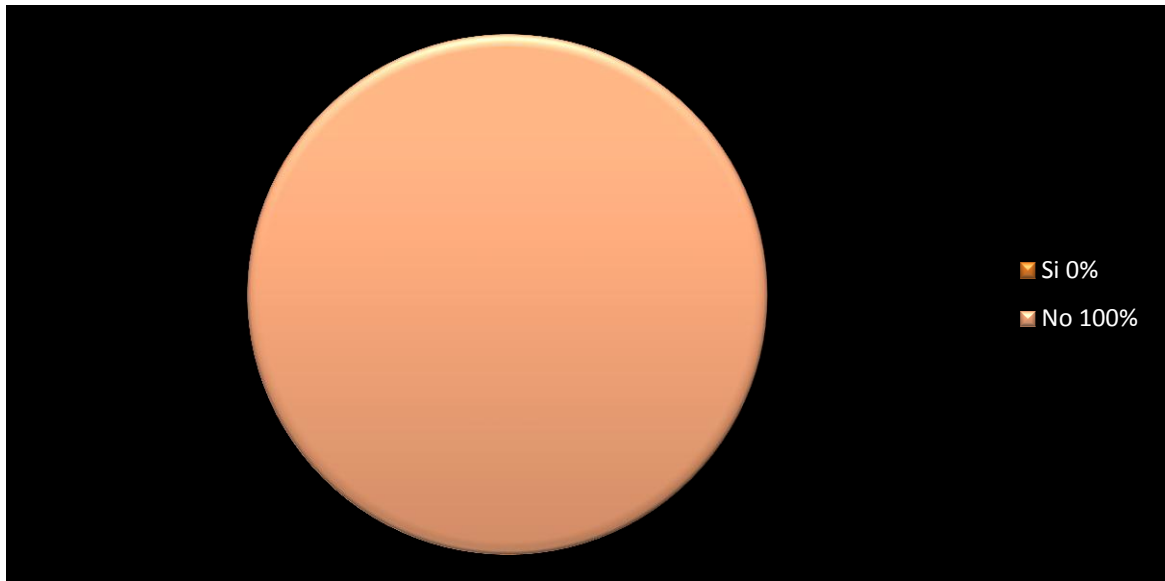
segun la Figura 26, entre los encuestados con 84% que corresponden a 16 personas consideran necesario contar con un mayor espacio para el numero promedio de estudianes que acceden a este para realizar las prácticas.

- **CONCLUSIÓN PARCIAL**

Se debe contar con espacio adecuado para que las prácticas puedan ser desplegada en total normalidad, permitiendo cada estudiante tener su propio equipo y que el docente pueda desplazarse sin ningun inconveniente.

13.1.8 ¿Los laboratorios cuentan con modalidades de acceso para estudiantes en condición de discapacidad?

Figura 27. Pregunta #8 desarrollada en la encuesta general



Fuente: Los autores

- **INTERPRETACIÓN**

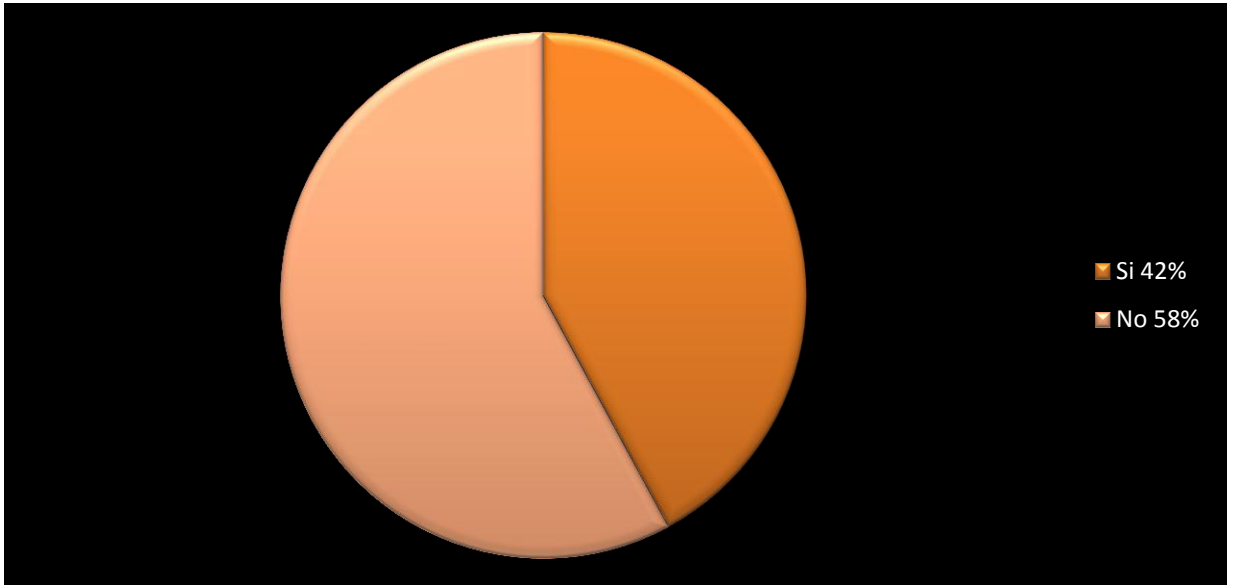
Segun la Figura 27, entre los docentes encuestados, con 100%, respondieron que no cuentan con este acceso a los laboratorios.

- **CONCLUSIÓN PARCIAL**

En el plan de mejoramiento se requiere contar con acceso para estudiantes con condicion de discapacidad.

13.1.9 ¿Los laboratorios cuentan con la posibilidad de instalar, desinstalar, y configurar software de acuerdo a necesidades de las asignaturas?

Figura 28. Pregunta #9 desarrollada en la encuesta general



Fuente: Los autores

- **INTERPRETACIÓN**

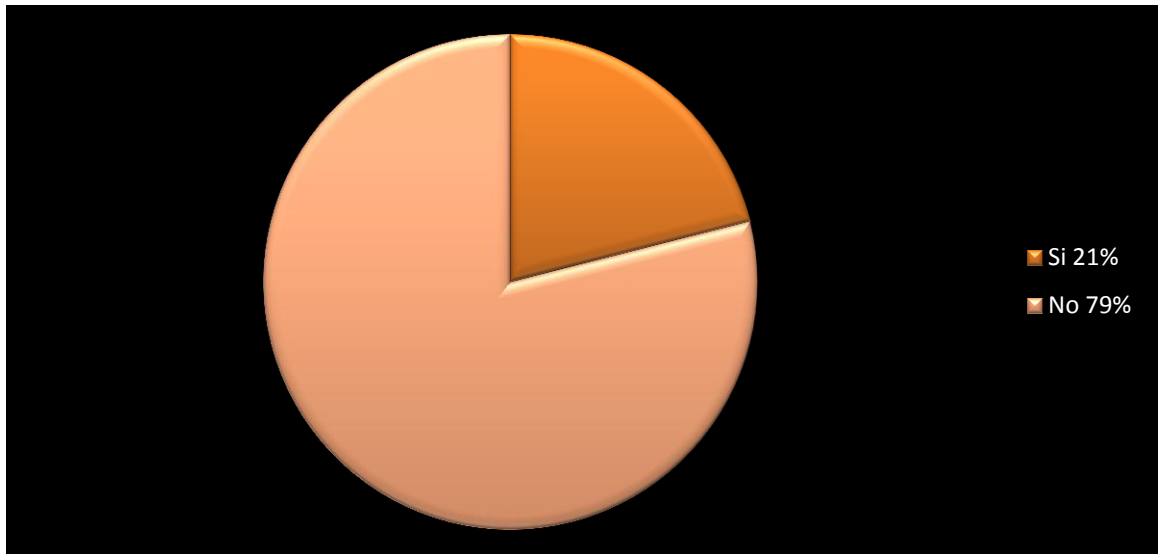
según la Figura 28, un 58% de los encuestados que corresponden a 11 personas argumentan que no cuentan con esta opción y requieren de poder administrar software en sus respectivas prácticas.

- **CONCLUSIÓN PARCIAL**

se debe contar con la posibilidad de realizar cualquier tipo de configuración de software, con el fin de brindarle todos los componentes necesarios que se deben tener en base a las asignaturas.

13.1.10 ¿Considera que el hardware de la universidad satisface las necesidades de las asignaturas tratadas en esta encuesta?

Figura 29. Pregunta #10 desarrollada en la encuesta general



Fuente: Los autores.

• INTERPRETACIÓN

Segun la Figura 29, entre los docentes encuestados con 79% que corresponden a 15 personas consideran que el hadware no cumple con las necesidades de las asignaturas, y un 21% que corresponden a 4 personas consideran que si satisfacen las necesidades de las asignaturas.

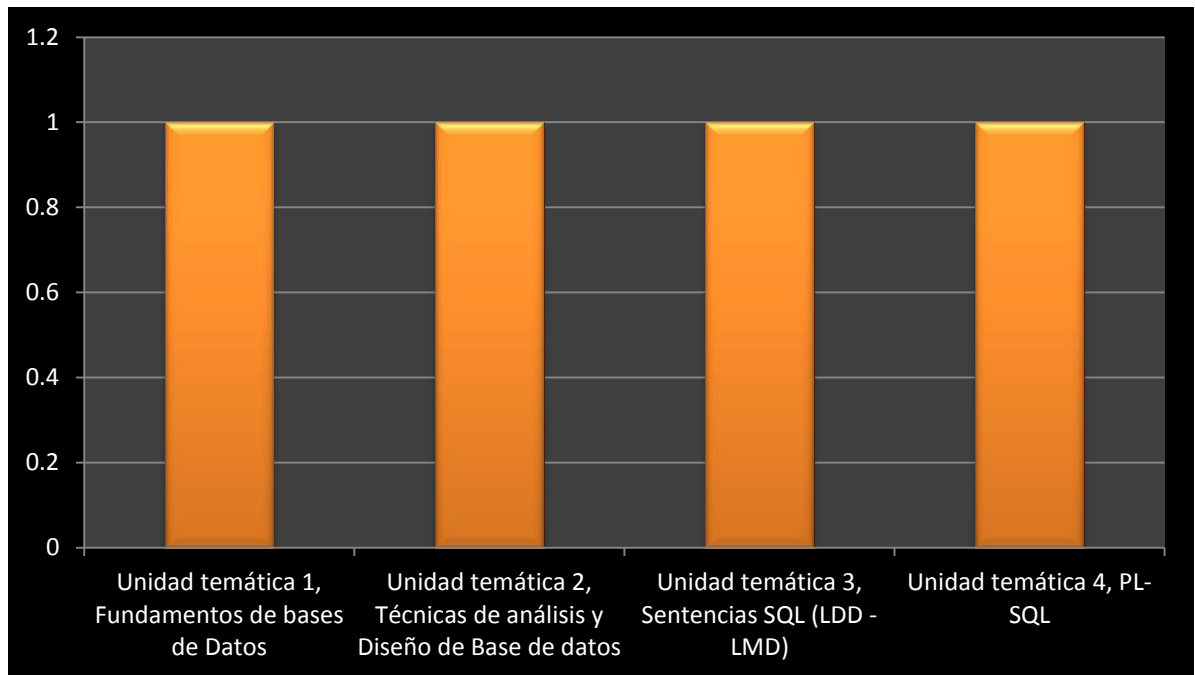
• CONCLUSIÓN PARCIAL

Se debe contar con el hardware necesario para todas las asignaturas con el fin de poder brindar una practica de calidad y acorde con lo requerido en cada una de ellas.

13.2 ENCUESTA ESPECIFICA REALIZADA A LOS DOCENTES A FINES CON EL LABORATORIO DE BASES DE DATOS

13.2.1 ¿Para cuáles de las unidades temáticas se presenta dificultad al realizar las prácticas propuestas en el currículo de la materia?

Figura 30. Pregunta #1 desarrollada en la encuesta específica



Fuente: Los autores

• INTERPRETACIÓN

Según los datos de la Figura 30, se puede analizar que en todas las unidades se presentan dificultades por igual en todas las temáticas al desarrollar las prácticas.

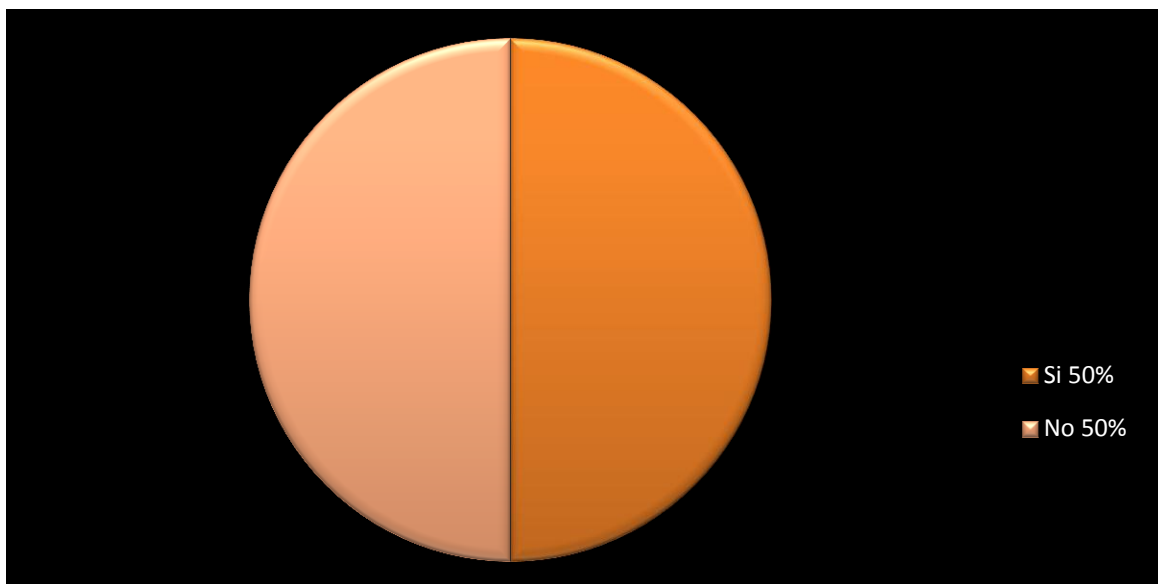
• CONCLUSIÓN PARCIAL

Se debe mejorar el desarrollo de las unidades y proponer nuevos esquemas para solucionar las dificultades presentadas en la asignatura de fundamentos de bases de datos.

UNIDAD TEMÁTICA NO.2: TÉCNICAS DE ANÁLISIS Y DISEÑO DE BASE DE DATOS

13.2.2 ¿Cree usted que el software actual de modelado de datos cumple con el propósito de las prácticas?

Figura 31. Pregunta #2 desarrollada en la encuesta específica



Fuente: Los autores

• INTERPRETACIÓN

Segun la Figura 31, un 50% de los encuestados argumentan que el software actual de modelado de datos cumple con los propositos de las prácticas y un 50% no esta satisfecho con el software.

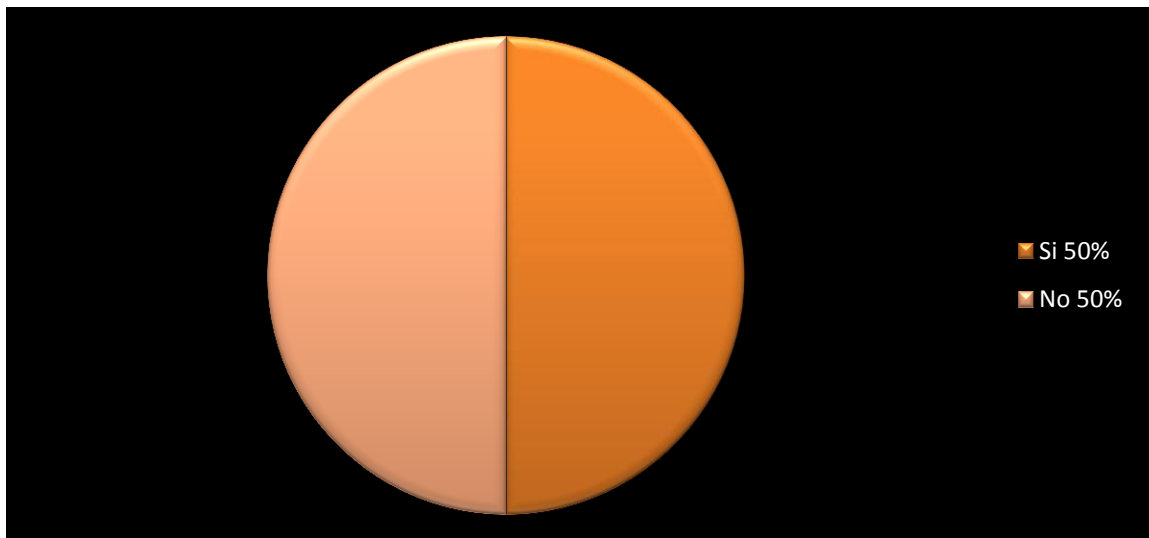
• CONCLUSIÓN PARCIAL

Se debe contar con el software correspondiente según el docente para poder cumplir los objetivos en las prácticas.

UNIDAD TEMÁTICA NO.3: SENTENCIAS SQL (LDD - LMD)

13.2.3 ¿Considera usted que el software actual de gestor de base de datos cumple con el propósito de las prácticas?

Figura 32. Pregunta #3 desarrollada en la encuesta específica



Fuente: Los autores

- **INTERPRETACIÓN**

Segun la Figura 32, la mitad de los encuestados considera que el software actual de gestor de bases de dastos cumple con los propositos de las practias y la parte restante considera lo contrario.

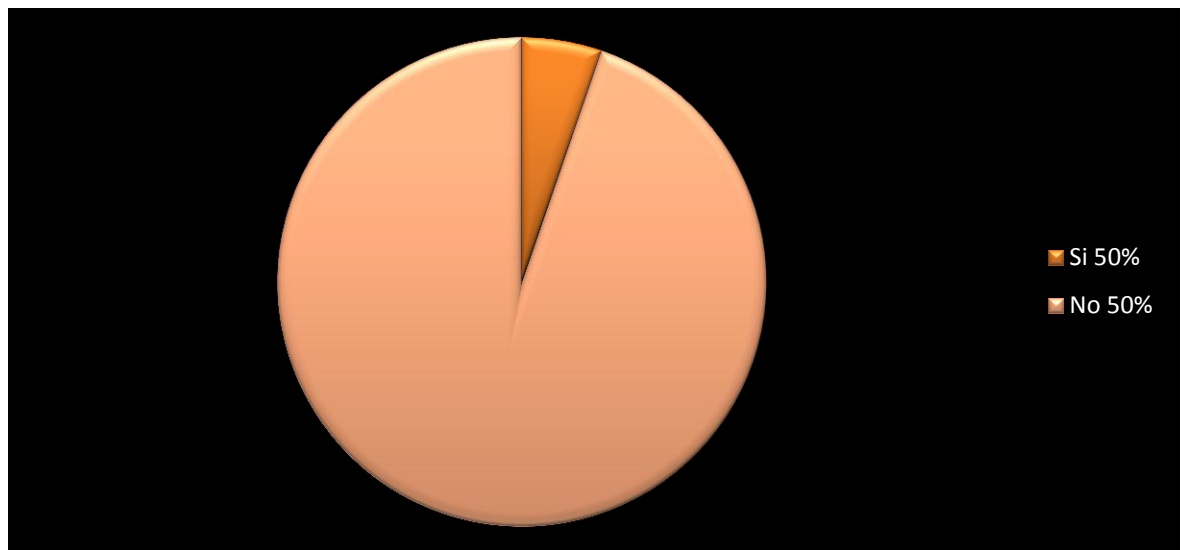
- **CONCLUSIÓN PARCIAL**

Es necesario que los estudiantes y profesores puedan contar con software de gestor de bases de datos requerido para poder cumplir con el proposito de las prácticas.

UNIDAD TEMÁTICA NO.4: PL-SQL

13.2.4 ¿Cree usted que el software actual, cumple con el propósito de las prácticas?

Figura 33. Pregunta #4 desarrollada en la encuesta específica



Fuente. Los autores

- **INTERPRETACIÓN**

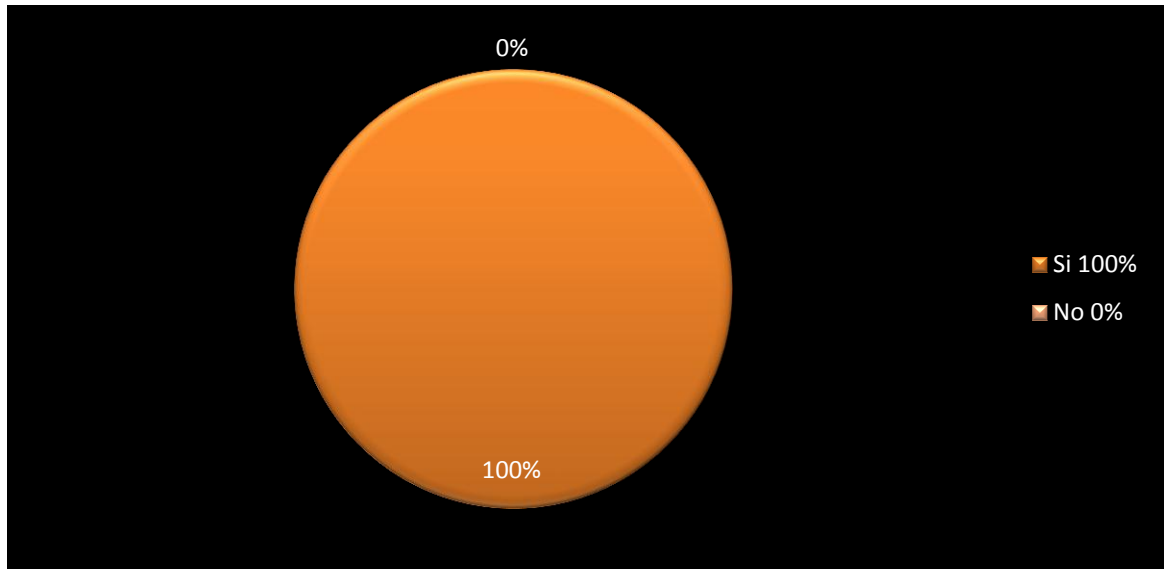
Según la Figura 33, el 50% de los encuestados consideran que el software actual cumple con los propósitos de las prácticas; el restante no lo perciben así.

- **CONCLUSIÓN PARCIAL**

Proporcionar el software adecuado a los docentes en la asignatura para poder cumplir con el propósito de las prácticas.

13.2.5 ¿Considera usted apropiado la implementación de tecnologías de virtualización en apoyo a las prácticas de laboratorio de bases de datos de forma remota?

Figura 34. Pregunta #5 desarrollada en la encuesta específica



Fuente: Los autores

- **INTERPRETACIÓN**

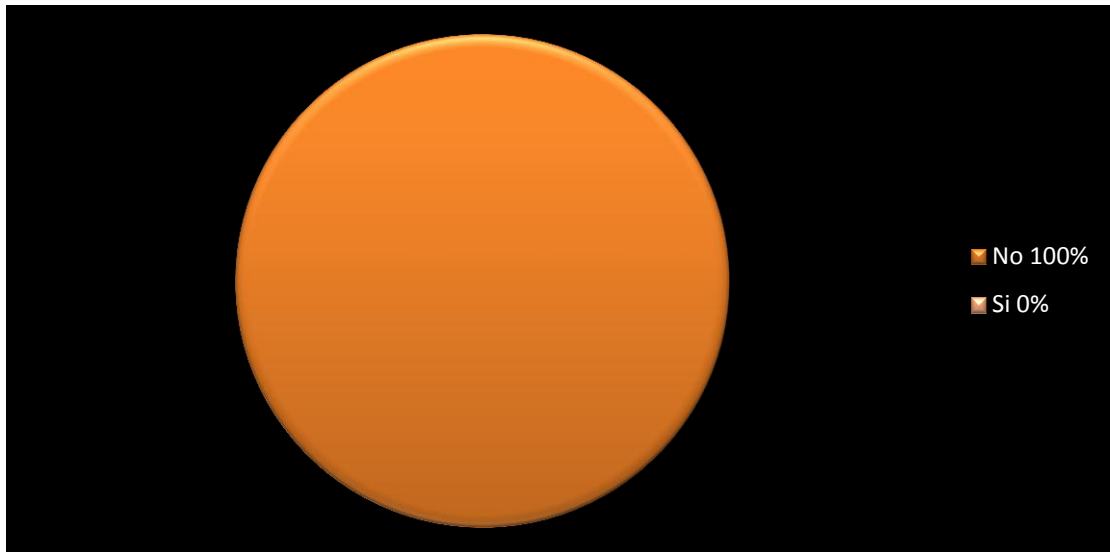
Segun la Figura 34, los encuestados ven con suma importancia, contar con tecnologías de virtualizacion, que apoyan el proposito de las prácticas.

- **CONCLUSIÓN PARCIAL**

Es de gran importancia que el laboratorio brinde tecnologías de virtualizacion como apoyo al proposito de las prácticas.

13.2.6 ¿ Con los recursos actuales que se cuenta en los laboratorios es posible que usted realice prácticas en big-data?

Figura 35. Pregunta #6 desarrollada en la encuesta específica



Fuente: Los autores.

- **INTERPRETACIÓN**

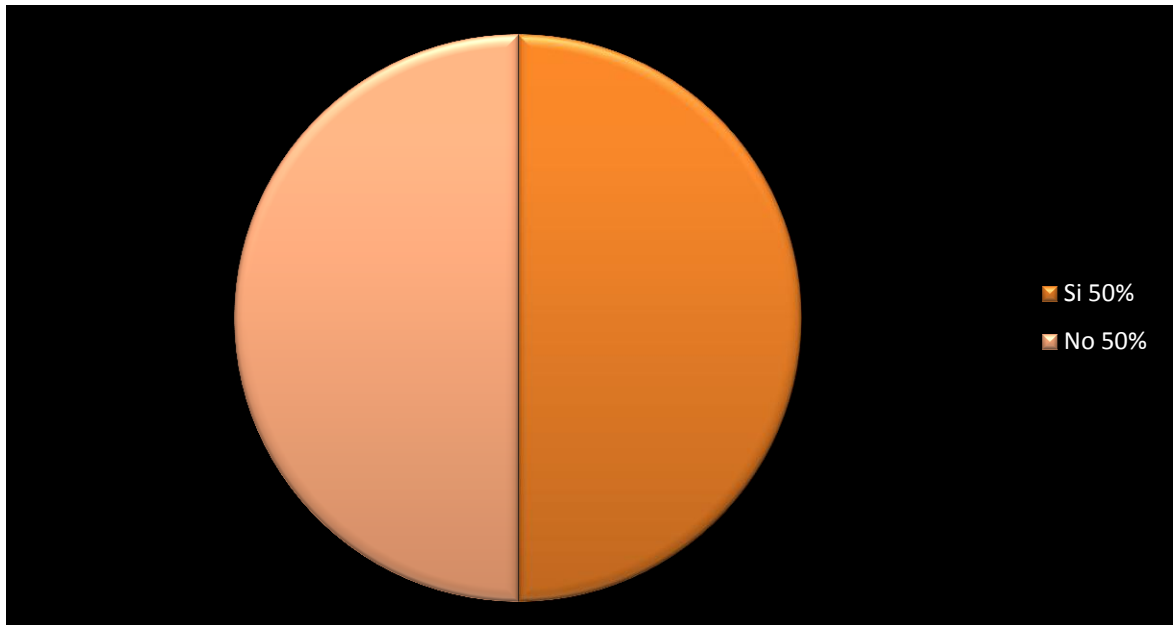
Según la Figura 35 y los encuestados, no disponen de recursos que posibiliten realizar prácticas en Big-Data.

- **CONCLUSIÓN PARCIAL**

Es necesario que los laboratorios brinden con recursos que faciliten las prácticas de Big-Data y de esta manera los estudiantes amplien sus conocimientos en esta tematica.

13.2.7 ¿Con los recursos actuales que se cuenta en los laboratorios es posible que usted realice prácticas en data warehouse?

Figura 36. Pregunta #7 desarrollada en la encuesta específica



Fuente: Los autores

• INTERPRETACIÓN

Segun la Figura 36, la mitad de los encuestados logran realizar prácticas en Data Warehouse pero no se cuentan con la infraestructura ideal para poder desplegarlas.

• CONCLUSIÓN PARCIAL

Se debe contar con infraestructura real para poder desarrollar prácticas optimas en Data Warehouse.

De acuerdo a los analisis obtenidos a partir de las encuestas general y especifica, teniendo en cuenta el estado actual de los laboratorios de la Universidad Catolica de Colombia, para el programa de ingenieria de sistemas, se despliega una serie de requerimientos descritos a continuacion en cada tabla.

13.3 ESPECIFICACIÓN DE REQUERIMIENTOS

El diseño propuesto en el siguiente capítulo, cuenta con características funcionales, que permite cumplir con el propósito de las prácticas. A continuación, se presenta el listado de requerimientos funcionales a tener en cuenta en el diseño del laboratorio de bases de datos.

Tabla 15. Requerimiento 1

REQUERIMIENTO FUNCIONAL	
ID	R01
Nombre	Instalar software requerido que permitan el desarrollo de las prácticas
Fecha	18/10/2014
Descripción	Es necesario que los estudiantes y profesores puedan instalar software requerido para poder desplegar las prácticas completamente.

Fuente: Los autores

Tabla 16. Requerimiento 2

REQUERIMIENTO FUNCIONAL	
ID	R02
Nombre	Contar con acceso remoto que apoye el desarrollo de prácticas dirigidas y libres
Fecha	18/10/2014
Descripción	Es de gran relevancia que el laboratorio de bases de datos cuente con acceso remoto, con el fin de apoyar al estudiante tanto en prácticas dirigidas como libres brindandoles espacios extras para resolver dudas.

Fuente: Los autores

Tabla 17. Requerimiento 3

REQUERIMIENTO FUNCIONAL	
ID	R03
Nombre	Contar con los permisos de usuarios necesarios para la realizacion de las prácticas.
Fecha	18/10/2014
Descripción	Es de gran importancia que los usuarios tanto estudiantes como profesores cuenten con permisos necesarios de usuarios para poder llevar acabo sus prácticas en su totalidad.

Fuente: Los autores

Tabla 18. Requerimiento 4

REQUERIMIENTO FUNCIONAL	
ID	R04
Nombre	Proporcionar teconologias de virtualizacion
Fecha	18/10/2014
Descripción	Se debe contar con componentes de virtualizacion con el fin de aprovechar los recursos de hadware para desarrollar mejores prácticas.

Fuente: Los autores

Tabla 19. Requerimiento 5

REQUERIMIENTO FUNCIONAL	
ID	R05
Nombre	Contar con mayor espacio para el número de estudiantes que acceden a realizar las prácticas
Fecha	18/10/2014
Descripción	Se debe contar con espacio adecuado para que las prácticas puedan ser desplegada en total normalidad, permitiendo cada estudiante tener su propio.

Fuente: Los autores

Tabla 20. Requerimiento 6

REQUERIMIENTO FUNCIONAL	
ID	R06
Nombre	Contar con modalidades de acceso para estudiantes en condición de discapacidad
Fecha	18/10/2014
Descripción	Se requiere contar con acceso a estudiantes con condición de discapacidad, con el fin de que puedan realizar sus prácticas sin ningún impedimento.

Fuente: Los autores

Tabla 21. Requerimiento 7

REQUERIMIENTO FUNCIONAL	
ID	R07
Nombre	Proporcionar configuración libre de software por el docente como por el estudiante.
Fecha	18/10/2014
Descripción	Se debe contar con la posibilidad de realizar cualquier tipo de configuración de software, con el fin de brindarle todos los componentes necesarios que se deben tener en base a las asignaturas.

Fuente: Los autores

Tabla 22. Requerimiento 8

REQUERIMIENTO FUNCIONAL	
ID	R08
Nombre	Satisfacer las necesidades de hardware en las asignaturas tratadas en esta encuesta
Fecha	18/10/2014
Descripción	Se debe contar con el hardware necesario para todas las asignaturas con el fin de poder brindar una practica de calidad.

Fuente: Los autores

Tabla 23. Requerimiento 9

REQUERIMIENTOS ESPECÍFICOS	
ID	RF09
Nombre	Brindar recursos para prácticas en Big-Data
Fecha	30/10/2014
Descripción	Es necesario que los laboratorios brinden con recursos que faciliten las prácticas de Big-Data y de esta manera los estudiantes amplien sus conocimientos en esta tematica.

Fuente: Los autores

Tabla 24. Requerimiento 10

REQUERIMIENTOS ESPECÍFICOS	
ID	R10
Nombre	Proveer infraestructura real para apoyar prácticas en Data Warehouse
Fecha	30/10/2014
Descripción	Se debe contar con recursos como infraestructura real para poder desarrollar prácticas optimas en Data Warehouse.

Fuente: Los autores

14. DISEÑO DEL LABORATORIO DE BASES DE DATOS

En este apartado se presenta el diseño del laboratorio de bases de datos y para el mismo, no se aplica una técnica o metodología especial para diseñarlo sino más bien, es dependiente de las necesidades del plan de estudios, es decir, es algo muy particular para cada programa pero que debe responder con los requerimientos y con el perfil de egreso de los estudiantes. Por su parte, aclarar que tampoco existe un modelo que abrigue completamente el laboratorio, simplemente, y tomando de referencia las necesidades, se ha elaborado un diseño particular del mismo. Este diseño tiene muy en cuenta la normatividad técnica asociada al laboratorio donde confluyen temas afines de protocolos, transmisión de datos, normas eléctricas, de redes y así sucesivamente.

Para iniciar el diseño se ha planteado un esquema conceptual que incorpora la infraestructura y los servicios del laboratorio. Inicialmente se expone el esquema de los servicios (ver Figura 37) del laboratorio donde se recogen los elementos académicos más importantes del mismo.

Figura 37. Esquema servicios laboratorio BD



Fuente: Los autores

14.1 SERVICIOS

A través del diseño, el laboratorio de base de datos ofrecerá servicios, que brindan beneficios a la comunidad estudiantil de la Universidad Católica de Colombia. Estos servicios son: los estudiantes podrán contar con permisos de usuarios e instalación del software requerido para poder desarrollar las prácticas, contar con acceso remoto a laboratorio con el fin de que los estudiantes puedan acceder desde cualquier lugar, sin tener que estar físicamente en el laboratorio, brinda tecnologías de virtualización donde a los estudiantes se les entregaran una máquina virtual con dos sistemas operativos: Linux y Windows, ofreciendo todos las aplicaciones requeridas por ellos, además podrán administrar recursos de hardware brindando una mayor flexibilidad y utilización de los mismos. Se ofrece herramientas para la realización de prácticas en BIGDATA, DATA WHAREHOUSE y DATAMINING; apoyado en las unidades de la asignatura de fundamentos de bases de datos.

A continuación, se describirán los componentes (ver Figura 38) que se tienen en cuenta para el desarrollo del diseño del laboratorio de bases de datos, para el programa de ingeniería de sistemas de la Universidad Católica de Colombia, detallando los elementos del laboratorio, tales como: adecuaciones físicas, diseño de la red, sistema de aire acondicionado, sistema de iluminación, sistema de seguridad y control, distribución de los equipos de cómputo, sistema cableado, etc. Debido a los constantes cambios y crecimiento a futuro este diseño será modular y escalable con el fin de adaptarse a los avances que pueden surgir en el transcurso del tiempo.

Figura 38. Componentes para el diseño del laboratorio BD



Fuente: Los autores.

El diseño del laboratorio está fundamentado en cuatro componentes, distribuidos en niveles, descritos a continuación:

- **Seguridad**, se definen las medidas preventivas a tener en cuenta para un óptimo funcionamiento tales como: prevención de incendios, desastres, rutas de evacuación, etc.

- **Hardware**, se establecen todos los dispositivos requeridos para el funcionamiento del laboratorio, tales como: equipos de cómputo, servidores, etc.
- **Infraestructura**, distribución de los espacios o lugares, donde estarán alojados los equipos del laboratorio, sillas y mesas, basados en normas de ergonomía y seguridad del trabajo.
- **Software**, las herramientas necesarias para el desarrollo de las prácticas, apoyadas en el plan curricular, tales como: herramientas de BigData, Data Mining, etc.

De acuerdo a las necesidades identificadas en los anteriores capítulos se establece que es de suma importancia contar con elementos como hardware y software, que soporten los procesos de apoyo a las prácticas; por esta razón se requiere de componentes tales como: un rack, dos servidores, un sistema de aire acondicionado de precisión, 25 computadores y un tablero interactivo utilizados para orientar las prácticas.

Como alternativa de mercado, a continuación se presenta una opción de equipamiento técnico seleccionada para satisfacer el diseño teórico anterior y que hace parte de una posible implementación.

14.2 HARDWARE

14.2.1 Racks. Dentro del espacio asignado para la administración de los diferentes servidores, se contara con un gabinete principal de administración (Rack) de 42 unidades, el Rack deberá cumplir con los estándares ANSI/EIA/RS-310-D, IEC 297-2, DIN41497 parte 1 y DIN41494 parte 7. Los gabinetes se podrán configurar según las necesidades, en su interior brindara una versatilidad para colocar en diferentes posiciones los bastidores para colocar los equipos, además contaran con paneles modulares laterales desmontables, los cuales facilitaran una mejor instalación de los equipos y al momento de alinearse se podrán quitar teniendo un mejor acceso a los demás gabinetes para una mejor administración. Para el laboratorio de bases de datos se seleccionó el gabinete explicado a continuación (ver Tabla 25).

Tabla 25. Rack del laboratorio de bases de datos

Gabinete 42UR Profundo	
Unidades	42
Estándares	ANSI/EIA/RS-310-D
Material de fabricación	Acero rolado en frío calibre 12, 14 y 16, pintura en polvo electrostática color negro
Dimensión externa	Ancho 600mm Profundo 960mm Alto 2055mm(con ruedas)
Dimensión interna	Ancho 19" Profundo 80mm Alto 42U
Puertas	<ul style="list-style-type: none">• Puerta delantera: Cristal templado con cerradura y dos llaves, rango de apertura 180°• Puerta trasera: Ventilada con cerradura y dos llaves, rango de apertura 180°

Tabla 25. (Continuación)

<p>Paneles</p>	<ul style="list-style-type: none">• Paneles laterales: Desmontables 1885x670mm con posibilidad de instalar cerradura• Paneles superior/inferior: Para la entrada de cableado	 <p>Diagrama de un rack de servidores con dimensiones: 2055mm de altura, 960mm de profundidad y 600mm de ancho. El rack tiene una puerta lateral desmontable y estantes internos.</p>
-----------------------	---	--

Fuente: Los autores.

14.2.2 Servidores. Para el laboratorio se eligen dos servidores, los cuales están descritos de acuerdo a sus características por medio de la Tabla 26. Un servidor con control de aplicaciones, que se darán uso en las salas y el segundo para la administración y funcionalidad de acceso remoto y virtualización.

Tabla 26. Servidor rack del laboratorio de bases de datos

Servidor en rack PowerEdge R730	
Procesador	Procesadores Intel® Xeon® E5-2600 v3 1.6GHz,15M Cache,6.40GT/s QPI,No Turbo,No HT,6C/6T (85W) Max Mem 1600MHz
Chipset	Chipset Intel serie C610
Memoria	Hasta 768 GB (24 ranuras DIMM): DDR4 de 4 GB/8 GB/16 GB/32 GB hasta 2133 MT/s
Controladoras Raid	Interno: PERC S130 Externo: PERC H830
Tipo de memoria DIMM y velocidad	2133MT/s RDIMMs
Capacidad de memoria	4GB RDIMM, 2133MT/s, Single Rank, x8 Data Width
Disco duro	1 TB 7.2K RPM SATA 6Gbps 3,5 pulgadas de conexión en caliente del disco duro
Tarjeta secundaria de red	Tarjeta secundaria de Broadcom 5720 QP 1Gb Red
fuentes de alimentación	Individual, de conexión en caliente Fuente de alimentación (1 + 0), 495 W
Costo (US\$)	COP\$7.710.580



Fuente: Los autores.

14.2.3 Acondicionamiento Térmico. Es importante que la sala donde están alojados los servidores y entre otros equipos, debe mantener la temperatura adecuada para la operación. El control y el mantenimiento de la calefacción, la ventilación, y el aire acondicionado (HVAC) son aspectos a tener en cuenta para una óptima operación.

14.2.3.1 Aire acondicionado de precisión. Debido al continuo funcionamiento y procesamiento de los equipos que genera disipación de calor, se requiere tener un nivel de temperatura estable con el fin de evitar cualquier tipo de daño o falla en estos. Razón por la cual es necesaria la adecuación de un aire acondicionado de precisión; basándose en los beneficios ambientales y ecológicos que proporciona (ver Tabla 27). Estos son:

Diseñados específicamente para satisfacer las necesidades de densas cargas electrónicas, las que generan un calor más seco que en el caso de enfriadores de ambiente convencionales, estando preparados además para dar enfriamiento durante todo el año, 24 horas al día. Los equipos electrónicos requieren niveles más precisos de humedad y calidad de aire que los que se consiguen con los sistemas de enfriamiento convencionales.⁴⁴

Tabla 27. Aire acondicionado de precisión del laboratorio de bases de datos

Aire acondicionado de precisión refrigerado por agua de unidades	
Tipo	Refrigerador de aire vapor
Modelo	54.4kw-102kw
Certificación	CE
Filtro	eu4
Voltaje funcionamiento	380/400 VAC
Compresor de la marca	Copeland
Precio	US\$ 2,000-10,000



Fuente: Los autores.

⁴⁴ Aires de Precisión Disponible. [En línea] :<<http://www.acpower.com.mx/Aires.html>>> [citado en 18 de octubre de 2014]

14.2.4 Equipos de cómputo. Se seleccionó el equipo descrito a continuación en la Tabla 28, que brinda las características necesarias para el desarrollo de actividades propias de laboratorio. Para cumplir con el objetivo del laboratorio se contara con 25 equipos, de los cuales 24 destinados para los estudiantes y el restante para el docente.

Tabla 28. Equipos de cómputo del laboratorio de bases de datos

Optiplex 9030	
Procesador	Cuarta generación del procesador Intel® Core™ i5-4590S (6MB Caché, 3.00 GHz con HD4600)
Sistema operativo	Windows® 7 Professional, 64-bit, Español (Incluye licencia y medio de Windows 8.1 Pro)
Teclado	Teclado Dell KB212-B USB silencioso, Español)
Mouse	Mouse óptico Dell MS111 USB
Memoria	4GB de Memoria DDR3L a 1600MHz (1 DIMM)
Tarjeta inalámbrica	Tarjeta Inalámbrica Intel® Wireless N 7260 doble banda 802.11a/g/n 2x2 + Bluetooth 4.0 LE
Pantalla	Pantalla WLED de 23" de Alta definición completa (Full HD) (1920 x 1080) con capa antirreflejo
Disco duro	Disco Duro SATA de 500GB 5400 RPM
Tarjeta de video	Gráficos Intel® HD 4400
Precio	COP\$2.358.942




Fuente: Los autores.

14.2.5 Tablero interactivo. Dentro del laboratorio como apoyo a las prácticas, se incluye un tablero interactivo (ver Tabla 29) el cual facilita a los docentes desarrollar presentaciones sin requerir de proyectores ni el uso de marcadores, generando interactividad entre docente- estudiante.

Tabla 29. Tablero interactivo del laboratorio de bases de datos

SMART Board X885	
Sistema operativo	Windows, Linux, Mac, Android
Puntos de contacto	10 puntos
Tamaño	199.4 cm x 135.6 cm x 16.5 cm
Resolución	8192 * 8192
Tipo de escritura	dedo o una pluma
Peso	27.2 kg
Precio	COP\$ 11,104,356



Fuente: Los autores

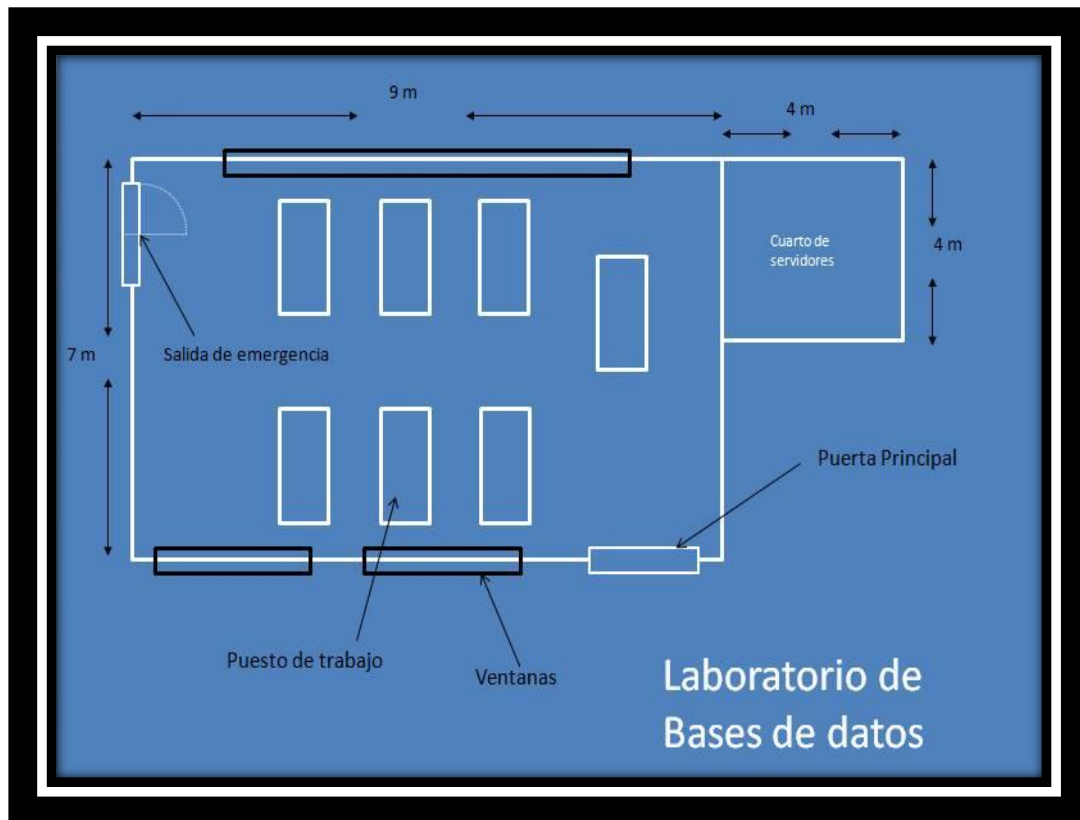
14.3 SOFTWARE

En este apartado, se incluyen las herramientas que fueron descritas en el marco conceptual por nombrar algunas: Avro, Zookeeper , Hive, Pig, Pentaho, Weka, entre otras; las cuales permiten el desarrollo de las prácticas en BigData, Datamining y Data Warehouse, siendo de suma importancia incorporarlas porque son nuevos esquemas vanguardistas para trabajar en las bases de datos.

14.4 INFRAESTRUCTURA

Este ítem del documento, tiene como objetivo diseñar las instalaciones teniendo en cuenta adecuaciones físicas tales como: iluminación, espacio de alojamiento de hardware, el área de servidores, área de sistema eléctrico, espacio trabajo y entre otras (ver Figura 39). Aclarando que esta es una aproximación del diseño del laboratorio de bases de datos fundamentado en normas que serán mencionadas en su momento.

Figura 39. Plano de distribución del laboratorio



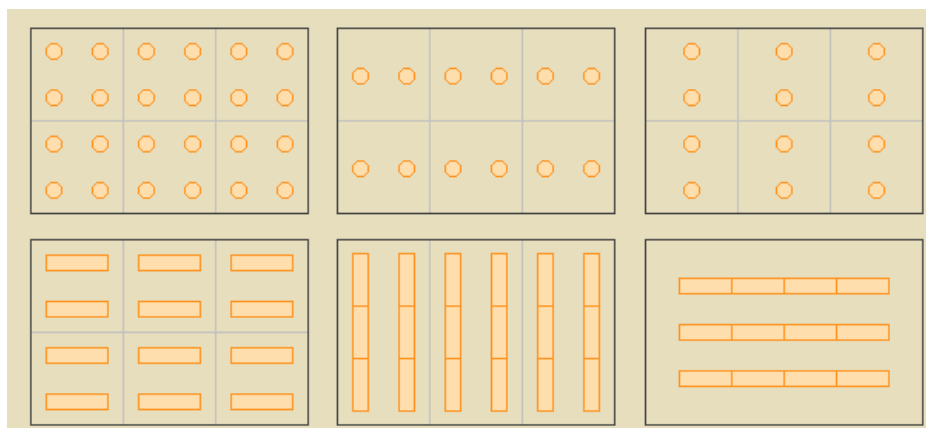
Fuente: Los autores

14.4.1 Adecuaciones físicas. Esta sección, tiene como propósito adecuar espacios físicos donde se alojaran los equipos de cómputo, servidores, racks y entre otros que hacen parte del funcionamiento del laboratorio de bases de datos.

14.4.1.1 Iluminación. El laboratorio debe contar con iluminación o alumbrado general de 350 a 1000 lux, el cual es utilizado en salas o laboratorios con propósitos de enseñanza y estudio, que a su vez tienen requisitos específicos debido al tipo de actividades que en estos se realizan. No contar con una óptima iluminación en las instalaciones del laboratorio puede generar fatiga visual, lesiones en la vista y no generar un ambiente confortable para el aprendizaje de los estudiantes.

“El alumbrado general proporciona una iluminación uniforme sobre toda el área iluminada. Es un método de iluminación muy extendido y se usa habitualmente en oficinas, centros de enseñanza, fábricas, comercios, etc. Se consigue distribuyendo las luminarias de forma regular por todo el techo.”⁴⁵

Tabla 30. Ejemplos de distribución de luminarias en alumbrado general



Fuente: <http://recursos.citcea.upc.edu/llum/interior/iluint1.html>

⁴⁵ Javier García Fernández. Iluminación de interiores[En línea] <<http://recursos.citcea.upc.edu/llum/interior/iluint1.html> >[Citado en 25 de Octubre de 2014]

14.4.1.2 Espacio de alojamiento de Hardware. Para el desarrollo de este proyecto es necesario contar con un espacio físico, el cual debe tener los equipos, sillas, mesas, servidores, racks; como también un espacio para el docente. Este lugar debe ser adecuado y acorde, teniendo en cuenta los espacios entre cada fila de computadores y el acceso a cada uno de ellos.

14.4.1.3 Sistema de distribución de energía. El sistema de distribución de energía incluye la energía principal en el laboratorio de bases de datos, los transformadores, el cableado, el sistema de tierra, tomacorrientes, y cualquier generador de energía, suministros de energía que se tiene para alimentar la energía al laboratorio.

- **Área del sistema eléctrico.**

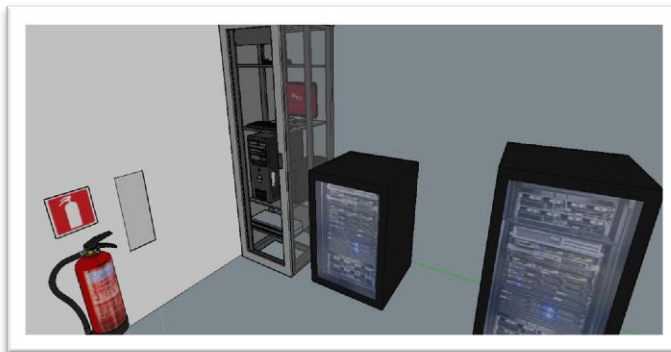
El cuarto eléctrico tiene que ver con el suministro de energía para todo el laboratorio, incluye paneles, conductores y tipos de receptores. Cabe resaltar que los voltajes pueden variar. También se destaca los sistemas de respaldo eléctrico.

- **Sistema de respaldo eléctrico.**

Es fundamental emplear medidas preventivas, con el fin de salvaguardar la integridad de la información, dado que este es el activo más importante dentro de una institución universitaria; por esta razón, se debe contar en el laboratorio de bases de datos con UPS, el cual es un componente crítico de alta disponibilidad, que en caso que la energía de la red eléctrica falle, la UPS suministrara energía 100% al hardware de acuerdo al tiempo o capacidad de suficiencia que esta disponga o en dado caso que se encienda algún generador de respaldo.

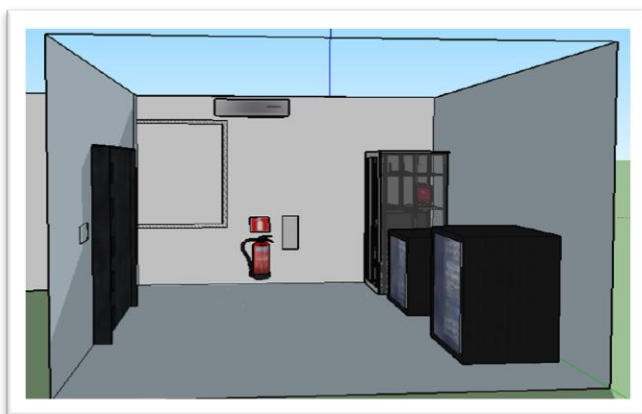
14.4.2 Área de Servidores. En este espacio se ubicaran los servidores (ver Figura 40 y Figura 41) los cuales permitirán el funcionamiento óptimo y adecuado del laboratorio de bases de datos; incluyendo un rack, donde se alojaran los dos servidores; uno destinado al acceso remoto, crear máquinas virtuales para los estudiantes en apoyo a sus prácticas y el otro encargado de administrar los recursos y redes del laboratorio. Además dentro de su estructura física, debe contar con una rampa (ver Figura 42) la cual permite un fácil acceso y salida de equipos; un piso falso, descrito a continuación.

Figura 40. Área de servidores



Fuente: Los autores

Figura 41. Área de servidores



Fuente: Los autores

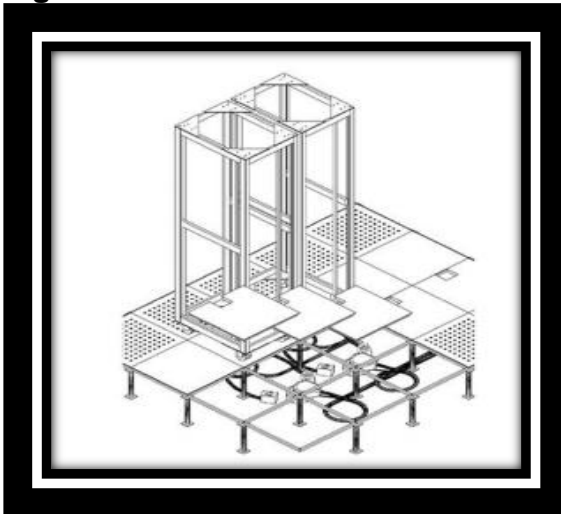
Figura 42. Rampa del área de servidores



Fuente: http://www.olaretta.com/index.php?option=com_content&view=article&id=62&Itemid=94&limitstart=3

14.4.3 Piso falso. Permite una cámara de aire entre el piso falso y el piso real, Facilita la distribución del mismo donde se requiera. Mejora la distribución y organización de las instalaciones. El propósito de un piso falso es canalizar el aire frío de las unidades de HVAC y dirigirlo hasta arriba para enfriar los equipos, encaminar los cables de la red y de la energía por debajo de este y actuar como un soporte seguro para la conexión a tierra de los equipos (ver Figura 43).

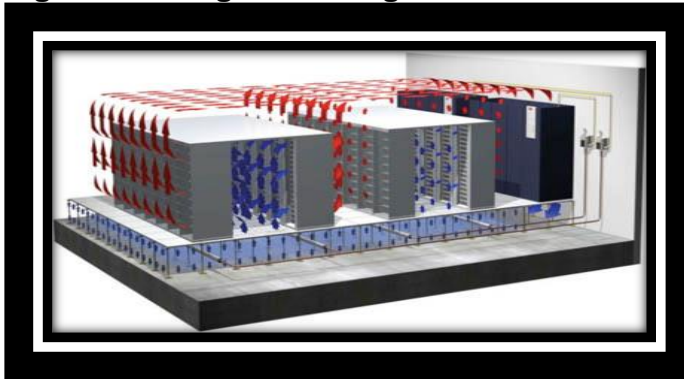
Figura 43. Piso falso



Fuente: <<http://dspace.ucuenca.edu.ec/jspui/bitstream/123456789/648/1/ts173.pdf>>

14.4.4 Acondicionamiento Térmico. Está dado por la adecuada distribución del aire a través del uso de un aire acondicionado de precisión; el cual en su parte superior toma el aire caliente y lo transforma en aire frío y este es canalizado por debajo del piso falso hacia los racks (ver Figura 44). Este método forma parte de la recomendación de la norma TIA-942.

Figura 44. Diagrama refrigeración



Fuente: <<http://www.mundohvacr.com.mx/mundo/2011/10/ahorro-energetico-en-sistemas-de-data-center/>>

14.4.5 Sistema de seguridad y control. El laboratorio de bases de datos debe contar con controles de seguridad física adecuados, basado en las normas AEN/CTN108 y ISO-27001, se describirá los siguientes componentes:

- Debe incluir detectores de movimientos, cámaras de vigilancia y alarmas para detectar intento de intrusión o robos.
- Contar con gráficos que indiquen el reglamento de uso de laboratorio y las restricciones en él.
- Tener luces y salidas de emergencia.
- Incluir dispositivos para la detención de agua con alarma en caso de inundaciones.

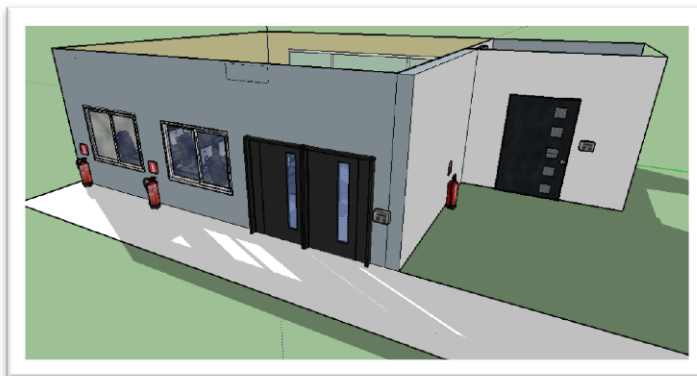
- Demarcar salidas de emergencia (ruta de evacuación).

14.4.5.1 Ambiental

- Contar con plan de mantenimiento preventivo a los equipos.
- Tener uso de cobertores para protección de polvo en los equipos.
- Contar con aire acondicionado adecuado para tener un buen confort en el laboratorio.

14.4.5.2 Sistema de detección de incendios y extinción. Debe ser regulado por la NFPA 70 y 75 para centros de cómputo; con el fin de salvaguardar la integridad física del laboratorio y de la información. El laboratorio contará con un sistema de alarma ante incendio y un sistema de apagado que no afecte el equipamiento electrónico. Este sistema incluye todos los dispositivos asociados con la detección de humo y extinción del fuego en el laboratorio de bases de datos (ver Figura 45).

Figura 45. Vista externa laboratorio BD y área de servidores



Fuente: Los autores.

14.4.5.3 Seguridad perimetral

- Se debe Incorporar sistemas de circuito cerrado de grabación de video para vigilancia del laboratorio.
- Utilizar barreras tales como: paredes y puertas de ingreso controlado para proteger áreas que contienen información y medios de procesamiento de la información.

Según las áreas de control de estándar UNE-ISO/IEC 17799 los requisitos mínimos que se deben tener en cuenta para la seguridad física del laboratorio son:

- **Área:**
 - las ventanas deben tener una malla que brinde seguridad y una película de recubrimiento oscuro para que la visualización sea interna, generando una protección externa del laboratorio.
 - Las estructuras de las paredes no deben tener ni permitir filtraciones externas, que puedan afectar el entorno.
 - Establecer reglas o parámetros de comportamiento para el ingreso al laboratorio como (prohibido comer, fumar, etc.).
 - El piso debe contar con un material antideslizante que permita prevenir accidentes del personal que labora.
 - Las dimensiones y características de puertas deberán cumplir con un ancho mínimo de 1,20 m para facilitar el acceso a personas con algún tipo de discapacidad.

- **Hardware:**
 - Para prever los desastres naturales como lo son los terremotos, es de suma importancia ubicar los equipos en sitios altos para evitar caídas.
 - Mantener la información y equipos de suma confidencialidad en áreas donde solo tenga acceso el personal admitido para su respectiva administración.
 - No permitir equipos móviles sobre los equipos con el fin de evitar caídas sobre ellos y que afecten su integridad.
 - Separar los equipos de las ventanas para evitar que caigan por ellas o que cualquier objeto lanzado del exterior los afecte.
 - Implementar pararrayos para evitar causar daños en los equipos por tormentas eléctricas.
 - Contar con un polo a tierra.
 - Realizar backup como medio de prevención y protección de la información y de esta manera evitar robos o sobrecargas de energías sobredimensionadas.

14.4.5.4 Sistema de control de acceso. El acceso al área de servidores debe ser estrictamente restringido y regulado, limitado únicamente al personal que lo administre, con el fin de evitar cualquier daños o accidentes sobre los equipos, cables o controles, puesto que hay equipos de hardware sensibles a daños. Por este motivo, para el ingreso del personal se utilizara un dispositivo biométrico para la identificación de la huella, el cual garantiza uno de los niveles de autenticación menos accesible en la actualidad.

14.4.5.5 Sistema de Cableado

- **Cableado estructurado**

El laboratorio debe contar con diferentes conexiones tecnológicas, utilizando un sistema de cableado estructurado siendo esta la forma más adecuada de interconectar varios y diferentes equipos de procesamiento de datos, además de que otorga una gran ventaja como es la de aceptar nuevas tecnologías soportando múltiples ambientes de computo. Cada una de las aplicaciones garantiza compatibilidad y calidad de acuerdo a los estándares ANSI/TIA/EIA-568^a- y ANSI/TIA/EIA-569.

- **Cableado de red**

El cableado dentro de un laboratorio de bases de datos se construye en la parte principal del flujo de la información, se debe garantizar la conectividad TCP/IP. La mayor parte de esto, se puede cumplir utilizando cables de categoría 5e, 6e o fibra óptica, esto depende de la funcionalidad de laboratorio, entender y conocer los requisitos del cableado, se convierte en una pieza fundamental al momento de construir el laboratorio.

14.5 ESPACIO DE TRABAJO

Se define como el espacio en el que los estudiantes ingresaran a realizar las prácticas del laboratorio; este lugar debe cumplir con una serie de características importantes para cumplir con el objetivo y estar adecuado para su uso.

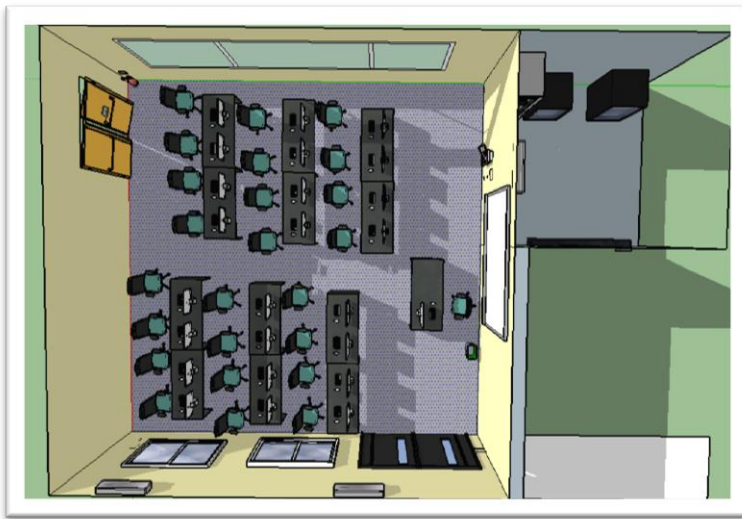
14.5.1 Espacio y movilidad. Con base al área actual de la sala 2, teniendo en cuenta el área individual del estudiante y el espacio que se debe contemplar entre equipos, las dimensiones utilizadas en este laboratorio son 9m x 7m, para un total de 63 m² de área, con una capacidad de 25 personas incluyendo al docente. (Ver Figura 46 y Figura 47).

Figura 46. Vista interna laboratorio BD



Fuente: Los autores

Figura 47. Vista desde arriba del laboratorio



Fuente: Los autores

Se plantea una distribución del espacio de trabajo de dos columnas, cada una con tres filas, donde cada una de ellas tendrá un total de 4 equipos de cómputo por fila de trabajo, pensado en la comodidad del estudiante y una óptima distribución del espacio de trabajo, además contará con un tablero inteligente el cual facilita presentaciones para el docente y el estudiante. (ver Figura 48 y Figura 49).

Figura 48. Distribución laboratorio BD



Fuente: Los autores

Figura 49. Distribución laboratorio BD



Fuente: Los autores

14.5.1.1 Puesto de trabajo. A su vez, debe existir un espacio reservado para las piernas permitiendo el confort de la postura de la persona en el sitio de trabajo, las dimensiones mínimas de los espacios libres para las piernas serán (ver Figura 50).

Figura 50. Distribución laboratorio BD



Fuente: Los autores

Tanto las sillas de trabajo como las mesas de trabajo, son de gran importancia para el confort de la persona, por lo cual deben satisfacer una serie de características de diseño; la silla deberá cumplir con las siguientes características:

- Regulable en altura (en posición sentado) margen ajuste entre 380 y 500 mm
- Anchura entre 400-450 mm
- Profundidad entre 380 y 420 mm
- Acolchado de 20 mm
- Borde anterior inclinado (gran radio de inclinación)

Espaldar:

- Anchura 400 - 450 mm
- Altura 250 – 300 mm

- Ajuste en altura de 150 – 250 mm

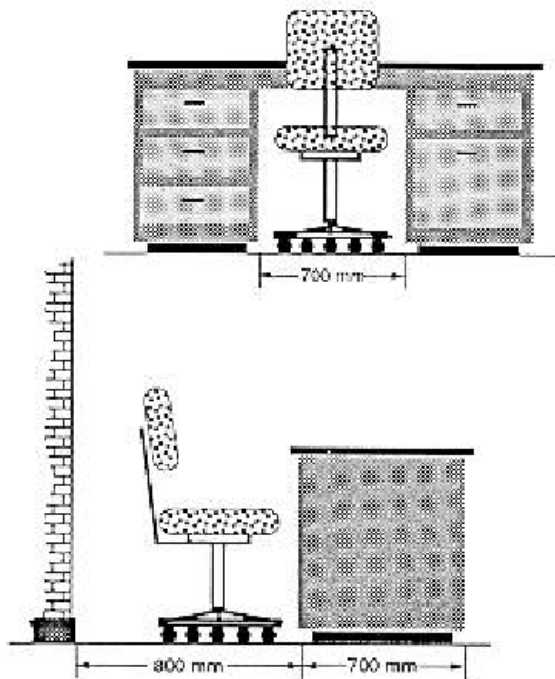
El espaldar alto deberá permitir el apoyo lumbar y ser regulable en inclinación.

- Regulación de la inclinación hacia atrás 15°.
- Anchura 300 - 350 mm.
- Altura 450 - 500 mm.
- Material igual al del asiento.

La mesa de trabajo deberá cumplir con las siguientes características:

- Si la altura es fija, ésta será de aproximadamente 700 mm.
- Si la altura es regulable, la amplitud de regulación estará entre 680 y 700 mm.
- La superficie mínima será de 1.200 mm de ancho y 800 mm de largo.
- El espesor no debe ser mayor de 30 mm.
- La superficie será de material mate y color claro suave, rechazándose las superficies brillantes y oscuras.
- Permitirá la colocación y los cambios de posición de las piernas.

Figura 51. Cotas de emplazamiento para las piernas en puestos de trabajo sentado



Fuente: http://www.insht.es/InshtWeb/Contenidos/Documentacion/FichasTecnicas/NTP/Ficheros/201a300/ntp_242.pdf

Es necesario que el plano de trabajo se sitúe a una altura adecuada a la talla del operario, ya sea en trabajos sentado o de pie. Dado que las medidas de las personas no son todas iguales, se deben cumplir con unos estándares generales para trabajos sentados (ver Figura 52).

Figura 52. Altura del plano de trabajo para puesto sentado (cotas en mm)

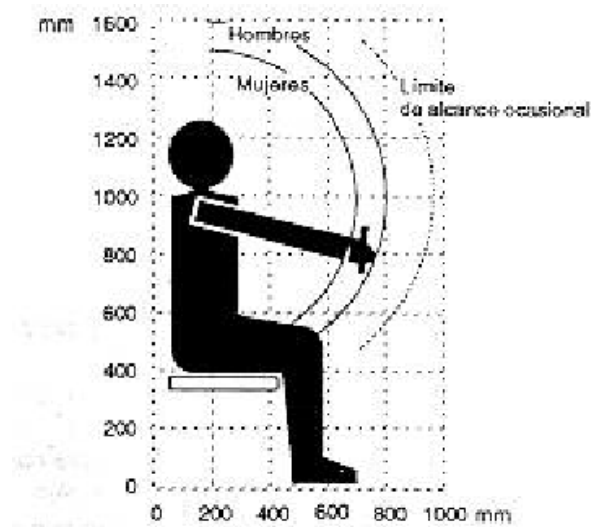


Fuente:

http://www.insht.es/InshtWeb/Contenidos/Documentacion/FichasTecnicas/NTP/Ficheros/201a300/ntp_242.pdf

Una buena disposición de los elementos de trabajo, no obligará a la persona a hacer movimientos forzados del tronco, evitando dolores de espalda. Por lo cual se deben determinar las distancias óptimas para obtener un confort en la postura adecuada como se observa en la Figura 53 y Figura 54.

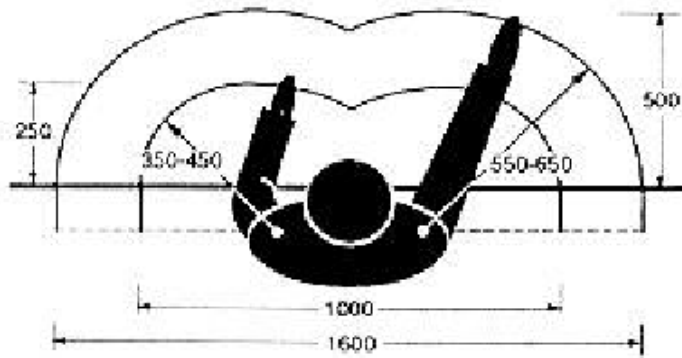
Figura 53. Arco de manipulación vertical en el plano sagital



Fuente:

http://www.insht.es/InshtWeb/Contenidos/Documentacion/FichasTecnicas/NTP/Ficheros/201a300/ntp_242.pdf

Figura 54. Arco horizontal de alcance del brazo y área de trabajo sobre una mesa



Fuente:

http://www.insht.es/InshtWeb/Contenidos/Documentacion/FichasTecnicas/NTP/Ficheros/201a300/ntp_242.pdf

15. VALIDACIÓN DEL DISEÑO DEL LABORATORIO

Este capítulo, tiene como propósito crear un esquema que represente los servicios que brinda el laboratorio junto con su infraestructura, validar que este diseño sea posible y apropiado para brindar apoyo a los propósitos académicos necesarios por las asignaturas tratadas anteriormente en el documento y acorde al levantamiento de requerimientos realizado. Sin embargo, es importante resaltar que no existe un proceso, normas o reglamentos que evalúen el diseño del laboratorio con enfoque en bases de datos.

De acuerdo con lo anterior, para la validación de este diseño de laboratorio, se utilizaron diversas herramientas, una de ellas fue la animación para analizar desde el punto de vista virtual como puede llegar hacer el laboratorio de bases de datos. Sin embargo desde el punto de vista teórico, se realizaron diferentes análisis sobre algunas herramientas que permiten validar ciertas características del laboratorio, una de ellas es utilizando herramientas relacionadas con la arquitectura empresarial de las organizaciones; en este sentido una de las más conocidas es TOGAF, la cual no es la única herramienta que existe pero si una de las más utilizadas y reconocidas en el diseño de arquitectura de software y así mismo en el de arquitecturas empresariales; por esta razón utilizamos el análisis de brecha de este framework para medir y evaluar los servicios prestados por el laboratorio.

Por otra parte, los servicios con los que cuenta el laboratorio, se observan en la Figura 55 sobre los cuales se realizara la respectiva validación del laboratorio.

Figura 55. Servicios del laboratorio de bases de datos



Fuente: Los autores.

15.1 VALIDACIÓN DE SERVICIOS

En este apartado, cada servicio es descrito y sustentado en una tabla de soporte en la cual se expone los componentes que proporciona el funcionamiento del servicio; seguido de un análisis de brecha, el cual realiza la comparación de los servicios actuales con los nuevos una vez finalizado el diseño del laboratorio de bases de datos.

15.1.1 Servicio de acceso remoto. El servicio de acceso remoto está diseñado para que los estudiantes puedan ingresar de forma remota, es decir sin estar físicamente en el laboratorio, con el fin de apoyar las prácticas sin estar dentro del laboratorio. Para poder contar con este servicio se implementó una infraestructura donde por medio de un servidor labs, los usuarios pueden interactuar y puedan iniciar una sesión segura. Luego de acceder el servidor de recurso brinda todos los recursos físicos y virtuales que tiene el laboratorio a través de este; el servidor de políticas en seguida de estar autenticados los usuarios, permite que cada usuario tenga acceso a los recurso con los permisos correspondientes y el servidor LDAP autentica que usuario puede o no acceder al portal e interactuar remotamente con el laboratorio. Este servicio esta soportado en (ver Tabla 31).

Tabla 31. Soporte para el servicio de acceso remoto

Infraestructura	Hardware	Software
Área de servidores	Servidor LDAP, LABS, aplicación de políticas, de recursos.	Windows server® 2012R2 Red Hat enterprise Linux 7
	Computadores, red, rack, máquinas virtuales	Vmware /VirtualBox

Fuente: Los autores.

15.1.2 Servicio de virtualización. El servicio de virtualización, que ofrece los laboratorios requiere de un mejor manejo y de unas características adicionales que permitan dar la funcionalidad y el enfoque correspondiente en apoyo a las prácticas. Este servicio está apoyado en (ver Tabla 32).

Tabla 32. Soporte para el servicio de virtualización

Infraestructura	Hardware	Software
Área de servidores	Servidor de virtualización	Windows server® 2012R2 Red Hat enterprise Linux 7
Espacio de trabajo	Computadores	Vmware /VirtualBox

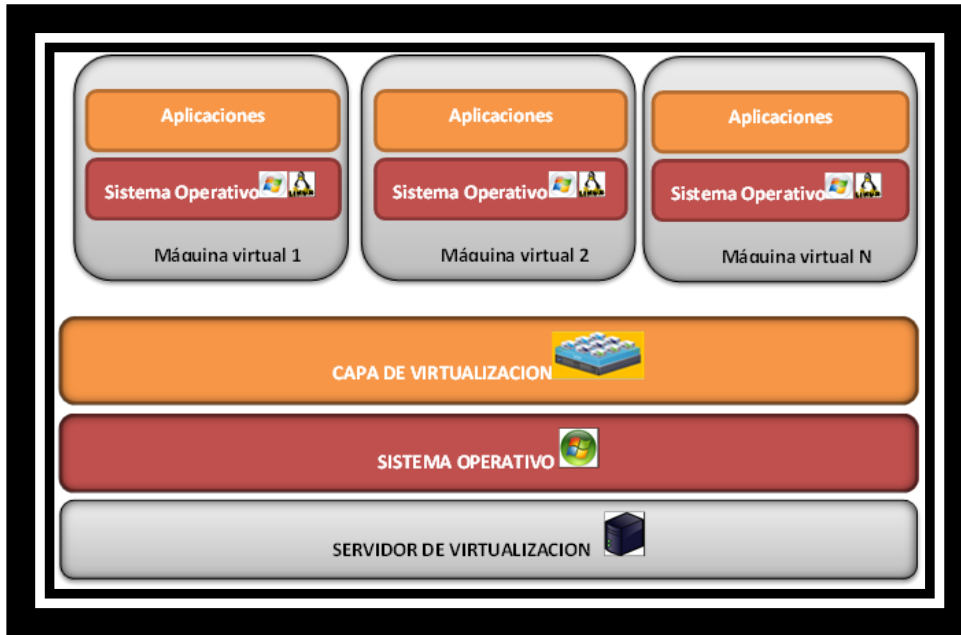
Fuente: Los autores.

Debido a que este servicio es apoyado a las prácticas de las asignaturas del área de administración de la información, es de resaltar que no es obligatorio estas características de software, estas varían dependiendo a las necesidades de los estudiantes en las prácticas, (ya sea Windows o Linux, utilizar Vmware o VirtualBox) aunque en el diseño se proponen las tecnologías más utilizadas en el mercado para que el usuario final decida cual es más acorde a su propósito o cual dese utilizar.

La virtualización según la tabla de soporte se contempla desde dos vistas diferentes, las cuales permiten la utilización del servicio desde el laboratorio como en equipos de cómputo externos.

En el esquema de virtualización (ver Figura 56) se observa como desde el servidor dedicado a la virtualización, proporciona “computadores o escritorios virtuales” los cuales pueden hacer uso de sistemas operativos Windows o Linux, los cuales cuentan con un usuario con todas las características de permisos necesarias para para poder instalar y desinstalar un sinfín de aplicaciones, de acuerdo a sus necesidades en las prácticas.

Figura 56. Esquema de servicio de virtualización



Fuente: Los autores.

15.1.3 Servicio de permisos administrativos a usuarios. Para contar con este servicio, se diseñó basado en las restricciones que se presentaban para desarrollar las prácticas, ya que los usuarios (estudiantes) no contaban con los mismos permisos administrativos. Para esto se establecieron permisos de acceso para el propósito en las prácticas, fundamentados también en la virtualización, permitiendo a los estudiantes la opción de instalar y configurar a su modo lo que se requiera para la práctica. Este servicio esta soportado en (ver Tabla 33).

Tabla 33. Soporte para el servicio de permisos administrativos a usuarios

Infraestructura	Hardware	Software
		Aplicaciones, sistema operativo
		Políticas de seguridad, permisos, restricciones

Fuente: Los autores.

15.1.4 Servicio de instalación de software. Para contar con este servicio, enfocado en el área de administración de la información, específicamente en la asignatura fundamento de bases de datos; se proporcionara al estudiante todo el software requerido para las prácticas, donde el laboratorio ofrecerá una máquina virtual con todas las herramientas solicitadas por el estudiante o en dado caso se tendrá instaladas en los computadores.

15.1.5 Servicio de componentes de enfoque específico. Para contar con este servicio, se diseñó en sustento a las necesidades de la materia fundamento de base de datos, brindando nuevas herramientas en apoyo a la realización de prácticas en BigData, Data Wharehouse y Datamining. Este servicio esta soportado en (ver Tabla 34).

Tabla 34. Soporte para el servicio de componentes de enfoque específico

Infraestructura	Hardware	Software
	computadores	Avro, Zookeeper
		Hive, Pig
		Pentaho Bi, Weka

Fuente: Los autores.

15.1.6 Servicio de práctica libre y dirigida. Para contar con este servicio, se estableció que el estudiante puede ingresar al laboratorio dentro de los horarios establecidos para realizar la práctica, donde se le asigna un equipo de acuerdo a las necesidades del usuario, también se puede proporcionar el software requerido al estudiante para que realice su práctica en su equipo y pueda desarrollarla estando en el laboratorio o accediendo remotamente desde cualquier lugar. La práctica dirigida la establece el docente estableciendo la hora y disponibilidad del laboratorio.

Como parte de la evaluación del diseño, se toma como base la matriz de análisis de brecha, la cual es tomada del capítulo 10.1 de TOGAF, en el framework la

matriz se utiliza para validar las arquitecturas de tecnología que se supone se brinda con el laboratorio. De acuerdo con la Tabla 35, el análisis de brecha permite comparar la arquitectura base con la arquitectura destino; la primera ubicada en la primera fila de forma vertical debe contener los servicios que el laboratorio brinda y la segunda ubicada horizontalmente en la primera fila de la tabla, debe contener los servicios que se obtendrá con el diseño del laboratorio y de esta manera identificar si cumple satisfactoriamente con las necesidades planteadas en los capítulos anteriores de este trabajo de grado.

De acuerdo con lo anteriormente mencionado es necesario tener en cuenta los siguientes aspectos:

- Si un servicio-funcionalidad esta tanto en la columna base como en la columna objetivo, se ubica un “incluido” en la intersección.
- Si un servicio- funcionalidad está en la columna base y no en la objetivo, se justifica por qué fue omitido, y si fue omitido este servicio por accidentalidad o porque realmente no es necesario, y es anotado en la columna “servicios eliminados”.
- Los servicios-funcionalidades que estén en la fila objetivo y no en la de base, se confirma que se encuentran en “Servicios nuevos”.

Tabla 35. Matriz análisis de brecha







Arquitectura de Destino Arquitectura base	Servicios							
	Servicio Virtualización	Servicio Acceso Remoto	Servicio Práctica Libre	Servicio Práctica Dirigida	Servicio Permisos a usuarios	Servicio de Instalación de software	Servicio Software específico	Servicios Específicos
Servicio Virtualización	Incluida (el servicio se mejoró para cumplir con los nuevos propósitos del laboratorio)							
Servicio Acceso Remoto								
Servicio Práctica Libre			Incluida					
Servicio Práctica Dirigida				Incluida				
Servicio permisos de Administrativos a usuarios					Incluida (se implementó el componente de virtualización para extender los permisos administrativos de usuarios)			
Servicio de Instalación de Software								
Servicio de Software específico							Incluida (se emplea software específico enfocado a fundamentos de bases de datos)	
Nuevo		Debe desarrollarse (servicio que se requiere para aumentar la disponibilidad del laboratorio)						Debe desarrollarse (el servicio que se requiere para que el estudiante pueda desplegar las prácticas a cabalidad, apoyados el servicio de virtualización)

Fuente: Los autores.

16. COSTEO DEL PROYECTO

A continuación se presenta los principales costos previstos en el desarrollo del diseño del laboratorio de bases de datos.

Tabla 36. Costeo del proyecto

Nombre	Descripción	Marca	Imagen	Costo (\$)	Cantidad
Servidor en rack	Procesador Procesadores Intel® Xeon® E5-2600v3 1.6GHz,15M Cache,6.40GT/s QPI,No Turbo,No HT,6C/6T (85W) Max Mem 1600MHz Chipset Chipset Intel serie C610Memoria Hasta 768 GB (24 ranuras DIMM): DDR4 de 4 GB/8 GB/16 GB/32 GB hasta 2133 MT/s	PowerEdge R730		15.421.000	2
Gabinete 42UR Profundo	Unidades: 42 Estándares ANSI/EIA/RS-310-D Material de fabricación Acero rolando en frío calibre 12, 14 y 16, pintura en polvo electrostática color negro / Dimensión externa Ancho 600mm Profundo 960mm / Alto 2055mm(con ruedas) Dimensión interna Ancho 19" / Profundo 80mm Alto 42U	Optronic		2.829.000	1
Aire acondicionado de precisión refrigerado por agua de unidades	Refrigerador de aire vapor/ Modelo 54.4kw-102kw Certificación CE /Filtro eu4 /Voltaje funcionamiento 380/400 VAC	Copeland		4.000.000	1
Switch	Se utilizan para la rápida interconexión de redes, sin embargo si cuentan con la función de Bridge, se ralentiza su funcionamiento, ya que se dedicarán a buscar errores en la información, pero en cuestiones de seguridad es mejor que utilicen en tal modo.	Dell		455.653	2
UPS	Smart-UPS Baterías intercambiables en caliente, sustituibles por el usuario doble conversion on line, regulacion de frecuencia y tension	Dell		5.200.000	2
Equipos de computo	Windows® 7 Professional, 64-bit, Español (Incluye licencia y medio de Windows 8.1 Pro)4GB de Memoria DDR3L a 1600MHz (1 DIMM)Pantalla WLED de 23" de Alta definición completa (Full HD) (1920 x 1080) con capa antirreflejo	optiplex 930		2.358.942	25
TOTAL:				107'955.676 \$	

Fuente: Los autores.

17. CONCLUSIONES Y APORTES

A continuación, se presentan las conclusiones y aportes relacionados con la investigación del diseño del laboratorio.

17.1 CONCLUSIONES

Este trabajo presentó el diseño del laboratorio de bases de datos de la universidad católica de Colombia para el programa de Ingeniería de Sistemas,

En este trabajo paralelamente se realizó un estado del arte de nuevos enfoques en bases de datos y nuevas formas de virtualización que pueden ser desplegados en el nuevo diseño del laboratorio.

Adicionalmente a los resultados mencionados, se realizó la revisión del estado actual de los laboratorios de ingeniería de sistemas en la Universidad Católica de Colombia, identificando cuatro problemáticas de carencia en el servicio tales como: no contar con acceso remoto al laboratorio y acceso a personas con condición de discapacidad, no contar con permisos de usuarios y con herramientas para el desarrollo de prácticas en BigData, Data Wharehouse y Datamining.

El diseño propuesto pretende brindar al usuario varios servicios tales como: poder acceder al laboratorio remotamente desde cualquier lugar sin estar físicamente en él, el estudiante con discapacidad parcial o permanente pueda ingresar a laboratorio sin ninguna dificultad, poder contar con servicios de usuarios para la realización de prácticas y contar con herramientas que apoyen el desarrollo de prácticas enfocadas en las temáticas de la asignatura de fundamento de bases de datos.

El diseño se realizó partiendo del análisis de requerimientos basados en las necesidades donde se detectaron 10 requerimientos que aportan en el ámbito de servicios y a su vez al mejoramiento del laboratorio.

Este diseño de laboratorio incluyó diferentes servicios nombrados anteriormente de los cuales se destaca el acceso remoto, puesto que era uno de los grandes inconvenientes con los que se enfrentaba el estudiante al verse limitado a realizar su práctica dentro del laboratorio; también es importante resaltar el servicio de virtualización como una solución a las restricciones que se presentaban en los permisos de usuario.

Por otra parte se presenta, el análisis de costeo que es una aproximación a la realidad de lo que puede costar el laboratorio para la Universidad

17.2 APORTES

En esta investigación enmarcada en el ámbito de la ingeniería de diseño, concretamente relacionado al análisis y diseño del laboratorio de base de datos, tiene como principales aportaciones las siguientes:

- Análisis y definición de requerimientos para el diseño del laboratorio de base de datos.
- La creación de un esquema basado en componentes, tomando como referencia para la elaboración del diseño.
- El diseño del laboratorio de bases de datos para la Universidad Católica de Colombia.
- La validación del diseño del laboratorio enmarcado en los servicios ofrecidos y basado en el análisis de brechas del framework Togaf.

18. BIBLIOGRAFÍA

- ADC Telecommunications, I. (2005). *Como diseñar un centro de datos óptimo*. Recuperado el 2014, de www.adc.com
- BARRERA, O. V. (1996). Recuperado el 2 de septiembre de 2014, de http://www.raco.cat/index.php/ensenanza/article/viewFile/21466/93439%3Forigin%3Dpublication_detail
- BARRIO, R., PARRONDO, J., BLANCO, E., & FERNANDEZ, J. (2011). Introducción de laboratorios virtuales en la enseñanza no presencial mediante entornos de trabajo propios. *Revista de Formación e Innovación Educativa Universitaria*, 55-67.
- cantabria, U. d. (2010). *Etapas del método hipotético-deductivo*. Recuperado el 05 de 08 de 2014, de [ocw.unican: http://ocw.unican.es/ciencias-de-la-salud/ciencias-psicosociales-i/materiales/bloque-i/tema-1/1.1.3.1-etapas-del-metodo-hipotetico-deductivo](http://ocw.unican.es/ciencias-de-la-salud/ciencias-psicosociales-i/materiales/bloque-i/tema-1/1.1.3.1-etapas-del-metodo-hipotetico-deductivo)
- GALLARDO GUTIERREZ, J. D. (2006). *MySQL 5*. Anaya.
- HAMACHER, Carl. (s.f.). *Organizacion de computadores*. S.A. MCGRAW-HILL / INTERAMERICANA DE ESPAÑA.
- IBM. (2014). *IBM*. Obtenido de <http://www.ibm.com/smarterplanet/us/en/smartbusiness/virtualdesktop>
- IMCYC. (Diciembre de 2006). *imcyc*. Recuperado el 30 de Julio de 2014, de <http://www.imcyc.com/revistact06/dic06/INGENIERIA.pdf>
- ISIDRO, Calvo. ZULUETA, Ekaitz. GANGOITI, Unai. LOPEZ, Jose. (s.f.). Laboratorios remotos y virtuales en enseñanzas técnicas y científicas. 1-21. Obtenido de <http://www.ehu.es/>
- javeriana.edu.co*. (s.f.). Recuperado el 13 de 08 de 2014, de http://portal.javeriana.edu.co/portal/page/portal/Facultad%20de%20Ingenieria/dpto_sist_laboratorios
- KONDABATHINI, V., BOUTAMINA, S., & VINJARAPU, S. K. (2011). A Theme to Unite The Resources of Different. *International Conference on Technology for Education*, 51-55.

- mcgraw-hill*. (s.f.). Recuperado el 30 de Julio de 2014, de <http://www.mcgraw-hill.es/bcv/guide/capitulo/8448148797.pdf>
- MEDINA, A. P., SABA, G. H., SILVA, J. H., & DURÁN, E. L. (2011). Los Laboratorios Virtuales y Laboratorios Remotos en la Enseñanza de la Ingeniería. *Revista Internacional De Educación En Ingeniería*.
- Microsoft. (2014). *Microsoft*. Obtenido de <http://www.microsoft.com/es-es/windows/enterprise/products-and-technologies/virtualization/operating-system/default.aspx>
- MUSA ZAMORA, R. (2012). Laboratorios Remotos: Actualidad y Tendencias Futuras. 113-118.
- NOVOA, N., & FLÓREZ, H. (2011). Los laboratorios virtuales adaptativos y personalizados en la educación superior. *REVISTA VÍNCULOS*.
- Portafolio.co*. (11 de Enero de 2013). Recuperado el 25 de 07 de 2014, de <http://www.portafolio.co/economia/presupuesto-el-pensamientoen-colombia>
- ROSA, Albert, FEISEL, Luyle. (2005). The Role of the Laboratory in Undergraduate Engineering Education. *Journal of Engineering Education*, 121-130.
- SICILIA, M. (28 de Septiembre de 2003). *cnx*. Recuperado el 30 de Julio de 2014, de <http://cnx.org/content/m17423/latest/>
- TAJANA SIMUNIC ROSING. (s.f.). *UCSDCSE Computer Science and Engineering*. Obtenido de <http://cseweb.ucsd.edu/classes/wi13/cse237A-a/handouts/introduction.pdf>
- Universidad de Cantabria. (s.f.). *Etapas del metodo hipotetico-deductivo*. Obtenido de <http://ocw.unican.es/ciencias-de-la-salud/ciencias-psicosociales-i/materiales/bloque-i/tema-1/1.1.3.1-etapas-del-metodo-hipotetico-deductivo>
- VirtualBox. (2014). Obtenido de <https://www.virtualbox.org/>

19. ANEXOS

Anexo A. Encuesta de carácter general para levantamiento de requerimientos del diseño de laboratorio

Este formulario tiene como objetivo realizar el levantamiento de requerimientos para la realización del diseño de laboratorios con los énfasis de arquitectura empresarial, bases de datos, arquitectura del computador y sistemas embebidos.

1. ¿CUALES MATERIAS, SEGUN SU ENFOQUE ACADEMICO, REQUIEREN DE UN INCREMENTO EN LA INFRAESTRUCTURA DE TI, PARA COMPLETAR EL DESARROLLO DE LAS PRÁCTICAS ACTUALES?

Selecciona todos los que correspondan.

- Arquitectura del computador
- Sistemas Operativos
- Digitales I y Laboratorio
- Digitales II y Laboratorio
- Sistemas Embebidos y Laboratorio
- Fundamento de Bases de Datos

2. ¿CREE USTED QUE LOS LABORATORIOS REQUIERE DE ACCESOS REMOTOS PARA LA REALIZACION DE LAS PRÁCTICAS?

Marca solo un ovalo.

Si

No

3. ¿EN LOS LABORATORIOS SE CUENTA CON LA POSIBILIDAD DE INSTALAR SOFTWARE QUE PERMITAN EL DESARROLLO DE LAS PRÁCTICAS EN DIFERENTES AMBIENTES?

Marca solo un ovalo.

Si

No

4. ¿LOS LABORATORIOS CUENTAN CON COMPONENTES DE ACCESO REMOTO QUE AMPLIEN LA DISPONIBILIDAD, Y APOYEN EL DESARROLLO DE PRÁCTICAS TANTO DIRIGIDAS COMO LIBRES?

Marca solo un ovalo.

Si

No

5. ¿CUENTA CON LOS PERMISOS DE USUARIO NECESARIOS PARA LA REALIZACION DE LAS PRÁCTICAS EN SU TOTALIDAD?

Marca solo un ovalo.

Si

No

6. ¿LOS LABORATORIOS CUENTAN CON EL COMPONENTE DE VIRTUALIZACION?

Marca solo un ovalo.

Si

No

¿Porque?

7. TENIENDO PRESENTE LA DISPOSICION ACTUAL DE LAS SALAS, BAJO SU CRITERIO ¿CREE QUE SE NECESITA UN MAYOR ESPACIO PARA EL NUMERO PROMEDIO DE ESTUDIANTES QUE ACCEDEN A ESTE PARA REALIZAR LAS PRÁCTICAS?

Marca solo un ovalo.

Si

No

¿Porque?

8. ¿LOS LABORATORIOS CUENTAN CON MODALIDADES DE ACCESO PARA ESTUDIANTES EN CONDICION DE DISCAPACIDAD?

Marca solo un ovalo.

Si

No

9. ¿LOS LABORATORIOS CUENTAN CON LA POSIBILIDAD DE INSTALAR, DES INSTALAR, Y CONFIGURAR SOFTWARE DE ACUERDO A NECESIDADES DE LAS ASIGNATURAS?

Marca solo un ovalo.

Si

No

10. ¿CONSIDERA QUE EL HARDWARE DE LA UNIVERSIDAD SATISFACE LAS NECESIDADES DE LAS ASIGNATURAS TRATADAS EN ESTA ENCUESTA?

Marca solo un ovalo.

Si

No

13. ¿Porque?

Anexo B. Encuesta de carácter específico para levantamiento de requerimientos del diseño de laboratorio

FUNDAMENTOS DE BASES DE DATOS

1. ¿PARA CUALES DE LAS UNIDADES TEMÁTICAS SE PRESENTA DIFICULTAD AL REALIZAR LAS PRÁCTICAS PROPUESTAS EN EL CURRÍCULO DE LA MATERIA?

- Unidad temática 1, Fundamentos de bases de Datos
- Unidad temática 2, Técnicas de análisis y Diseño de Base de datos
- Unidad temática 3, Sentencias SQL (LDD - LMD)
- Unidad temática 4, PL-SQL

Unidad temática No.2: Técnicas de análisis y Diseño de Base de datos

Modelo de Datos, Modelo Entidad Relación, Modelo Relacional, Normalización

2. ¿CREE USTED QUE EL SOFTWARE ACTUAL DE MODELADO DE DATOS CUMPLE CON EL PROPÓSITO DE LAS PRÁCTICAS?

Marca solo un ovalo.

Si

No

Unidad temática No.3: Sentencias SQL (LDD - LMD)

Creación de Bases de Datos y las estructuras de relación(LDD), Tareas básicas de gestión sobre una Base de Datos, Álgebra Relacional, Lenguaje de Manipulación de datos (LMD)

3. ¿CONSIDERA USTED QUE EL SOFTWARE ACTUAL DE GESTOR DE BASE DE DATOS CUMPLE CON EL PROPÓSITO DE LAS PRÁCTICAS?

Marca solo un ovalo.

Si

No

Unidad temática No.4 : PL-SQL

Store Procedure (Procedimientos almacenados) Triggers (Disparadores)

4. ¿CREE USTED QUE EL SOFTWARE ACTUAL, CUMPLE CON EL PROPÓSITO DE LAS PRÁCTICAS?

Marca solo un ovalo.

Si

No

De acuerdo a la opción seleccionada, describa ¿Porque?

5. ¿CONSIDERA USTED APROPIADO LA IMPLEMENTACIÓN DE TECNOLOGÍAS DE VIRTUALIZACIÓN EN APOYO A LAS PRÁCTICAS DE LABORATORIO DE BASES DE DATOS DE FORMA REMOTA?

Marca solo un ovalo.

Si

No

¿Porque?

6. ¿CON LOS RECURSOS ACTUALES QUE SE CUENTA EN LOS LABORATORIOS ES POSIBLE QUE USTED REALICE PRÁCTICAS EN BIG-DATA?

Marca solo un ovalo.

Si

No

¿Si la respuesta anterior fue no, ¿qué recursos se requieren en hardware y software para realizar esta práctica ?

7. ¿CON LOS RECURSOS ACTUALES QUE SE CUENTA EN LOS LABORATORIOS ES POSIBLE QUE USTED REALICE PRÁCTICAS EN DATA WAREHOUSE?

Marca solo un ovalo.

Si

No

¿Si la respuesta anterior fue no, ¿qué recursos se requieren en hardware y software para realizar esta práctica ?
