

UNIVERSIDAD TÉCNICA DE AMBATO



**FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA E
INDUSTRIAL**

CENTRO DE ESTUDIOS DE POSGRADO

MAESTRÍA EN GESTIÓN DE BASES DE DATOS

TEMA:

**METODOLOGÍA PARA MIGRACIÓN DE DATOS QUE PERMITA
ASEGURAR Y CONSERVAR LA INTEGRIDAD Y CONSISTENCIA DE
LA INFORMACIÓN ADMINISTRADA POR LA EMPRESA VSYSTEMS**

Trabajo de investigación científica

**Previa a la obtención del Grado Académico de Magister en Gestión de Bases
de Datos.**

AUTORA: Ing. Susana Beatriz Caraguay Martínez

DIRECTOR: Ing. Mg. Franklin Oswaldo Mayorga Mayorga

Ambato

2012

Al Consejo de Postgrado de la UTA.

El tribunal receptor de la defensa del trabajo de investigación con el tema: “METODOLOGÍA PARA MIGRACIÓN DE DATOS QUE PERMITA ASEGURAR Y CONSERVAR LA INTEGRIDAD Y CONSISTENCIA DE LA INFORMACIÓN ADMINISTRADA POR LA EMPRESA VSYSTEMS”, presentando por: Ing. Susana Beatriz Caraguay Martínez y conformado por: Ing. Mg. Elsa Pilar Urrutia Urrutia, Ing. Mg. Marco Antonio Jurado Lozada, Ing. Mg. Víctor Manuel Pérez Rodríguez miembros del Tribunal, e Ing. Mg. Franklin Oswaldo Mayorga Mayorga Director del trabajo de investigación y presidido por: Ing. M.Sc. Oswaldo Eduardo Paredes Ochoa, Presidente del Tribunal, Ing. Juan Enrique Garcés Chávez Director del CEPOS – UTA, una vez escuchada la defensa oral, el tribunal aprueba y remite el trabajo de investigación para uso y custodia en las bibliotecas de la UTA.

Ing. M.Sc. Oswaldo Eduardo Paredes Ochoa
Presidente del Tribunal de Defensa

Ing. Mg. Juan Enrique Garcés Chávez
DIRECTOR CEPOS

Ing. Mg. Franklin Oswaldo Mayorga Mayorga
Director de trabajo de Investigación

Ing. Mg. Elsa Pilar Urrutia Urrutia
Miembro del Tribunal

Ing. Mg. Marco Antonio Jurado Lozada
Miembro del Tribunal

Ing. Mg. Víctor Manuel Pérez Rodríguez.
Miembro del Tribunal

AUTORÍA DE LA INVESTIGACIÓN

La responsabilidad de las opiniones, comentarios y críticas emitidas en el trabajo de investigación con el tema: “METODOLOGÍA PARA MIGRACIÓN DE DATOS QUE PERMITA ASEGURAR Y CONSERVAR LA INTEGRIDAD Y CONSISTENCIA DE LA INFORMACIÓN ADMINISTRADA POR LA EMPRESA VSYSTEMS”, nos corresponde exclusivamente a: Susana Beatriz Caraguay Martínez y Franklin Oswaldo Mayorga Mayorga, Director del trabajo de investigación; y el patrimonio intelectual del mismo a la Universidad Técnica de Ambato.

Ing. Susana Beatriz Caraguay M.
AUTORA

Ing. Mg. Franklin Oswaldo Mayorga M.
DIRECTOR

DERECHOS DE AUTOR

Autorizo a la Universidad Técnica de Ambato, para que haga de éste trabajo de investigación o parte de él un documento disponible para su lectura, consulta y procesos de investigación, según las normas de la Institución.

Cedo los derechos de mi trabajo de investigación, con fines de difusión pública, además apruebo la reproducción de ésta, dentro de las regulaciones de la Universidad.

Ing. Susana Beatriz Caraguay Martínez

AGRADECIMIENTO

A Dios por guiarme con su perfecta sabiduría a lo largo de mi vida, por todas las bendiciones que cada día me regala.

A mis padres por su apoyo incondicional en cada proyecto profesional y personal que emprendo.

A la Universidad Técnica de Ambato por darme la oportunidad de profesionalizarme y alcanzar ésta meta académica.

Al Ing. Mg. Franklin Mayorga Director del trabajo de investigación por haber guiado durante el desarrollo del presente trabajo de investigación.

A los miembros de la Empresa VSYSTEMS, por su apoyo y colaboración en el desarrollo del presente proyecto.

Ing. Susana Caraguay M.

DEDICATORIA

El presente trabajo de investigación lo dedico con mucho cariño a mis padres y hermanos, que son mi guía y mi apoyo en todo momento.

A todos mis amigos que me han apoyado e inspirado para continuar con mi desarrollo profesional.

Ing. Susana Caraguay M.

ÍNDICE GENERAL

PRELIMINARES	Pág.
AUTORÍA DE LA INVESTIGACIÓN	iii
DERECHOS DE AUTOR.....	iv
AGRADECIMIENTO	v
DEDICATORIA	vi
ÍNDICE GENERAL	vii
INDICE DE TABLAS	x
ÍNDICE DE GRÁFICOS	xii
RESUMEN	xiv
SUMMARY	xv
INTRODUCCIÓN	xvi
TEXTO	
CAPÍTULO I	1
1. EL PROBLEMA.....	1
1.1. Tema.....	1
1.2. Planteamiento del problema.....	1
1.2.1. Contextualización	1
1.2.2. Árbol de problemas.....	4
1.2.3. Análisis Crítico	5
1.2.4. Prognosis	6
1.2.5. Formulación del problema	7
1.2.6. Interrogantes de la investigación.....	7
1.2.7. Delimitación del objeto de investigación.....	8
1.3. Justificación	8
1.4. Objetivos.....	10
CAPÍTULO II.....	11
2. MARCO TEÓRICO.....	11
2.1. Antecedentes Investigativos	11
2.2. Fundamentaciones	12
2.2.1. Fundamentación Filosófica	12

2.2.2.	Fundamentación Tecnológica.....	12
2.2.3.	Fundamentación Legal	13
2.3.	Categorías Fundamentales.....	14
2.3.1.	Red de inclusión.....	14
2.3.2.	Constelación de ideas de la Variable Independiente	15
2.3.3.	Constelación de ideas de la Variable Dependiente	16
2.4.	Hipótesis	26
CAPÍTULO III.....		27
3. METODOLOGÍA.....		27
3.1.	Modalidades Básicas de Investigación	27
3.2.	Tipos o Niveles de investigación.....	28
3.3.	Población y Muestra.....	28
3.4.	Operacionalización de Variables	30
3.4.1.	Operacionalización de Variable Independiente	30
3.4.2.	Operacionalización de Variable Dependiente.....	31
3.5.	Técnicas e instrumentos	32
3.6.	Validez y confiabilidad	32
3.7.	Plan de recolección de la información.....	33
3.8.	Plan de procesamiento de la información	34
CAPITULO IV		35
4. ANÁLISIS E INTERPRETACIÓN DE RESULTADOS		35
4.1.	Análisis e interpretación.....	35
4.2.	Verificación de la hipótesis	47
4.2.1.	Modelo Estadístico.....	47
4.2.2.	Definición del nivel de significancia	48
4.2.3.	Especificación de la región de aceptación y rechazo	48
4.2.4.	Recolección de cálculos y datos estadísticos	48
4.2.5.	Cálculo del Chi cuadrado	49
CAPITULO V.....		51
5. CONCLUSIONES Y RECOMENDACIONES		51
5.1.	Conclusiones	51
5.2.	Recomendaciones.....	52
6. PROPUESTA		53

6.1.	Datos Informativos	53
6.2.	Antecedentes de la propuesta	53
6.3.	Justificación	54
6.4.	Objetivos	54
6.5.	Análisis de factibilidad	55
6.6.	Fundamentación Teórica	57
6.6.1.	Gestor de base de datos Oracle	57
6.6.2.	Gestor de base de datos PostgreSQL	61
6.6.3.	Herramientas de migración	66
6.6.4.	Metodología de migración	69
6.6.4.1.	Metodología de migración Strategy 7 Corporation	69
6.6.4.2.	Database migration methodology from oracle to sql-server using-ssma...71	
6.6.4.3.	Metodología de Migración de datos (Álvaro Alejandro Herrera)	72
6.6.4.4.	Data Migration Methodology (JosephHudicka, Dulcian Inc)	76
6.6.5.	Resumen de fases de las metodologías de migración	78
6.6.6.	Cuadro comparativo de metodologías de migración	80
6.6.7.	Análisis de metodologías de migración.....	81
6.6.8.	Metodología de migración de datos propuesta para la empresa VSYSTEMS ..81	
6.7.	Metodología. Modelo Operativo	103
6.7.1.	Modelo Operativo	103
6.7.2.	Apliación de la metodologia de migración de base de datos en la empresa VSYSTEMS	104
6.7.2.1.	PM1: Preparación del Entorno de migración.	104
6.7.2.2.	PM2: Ejecución de la migración.....	118
6.7.2.3.	PM3: Evaluación y control de calidad	124
6.8.	Administración	125
6.9.	Previsión de la evaluación	125
7.	CONCLUSIONES Y RECOMENDACIONES	126
7.1.	Conclusiones	126
7.2.	Recomendaciones.....	127
	MATERIALES DE REFERENCIA.....	128
	BIBLIOGRAFÍA	128
	ANEXOS	132

INDICE DE TABLAS

Tabla N° 1: Alternativa de software libre para base de datos	12
Tabla N° 2: Unidades de Observación	29
Tabla N° 3: Operacionalización de la variable independiente.	30
Tabla N° 4: Operacionalización de la Variable Dependiente	31
Tabla N° 5: Plan de recolección de la Información	33
Tabla N° 6: Tabulación origen del proceso de migración	35
Tabla N° 7: Tabulación complejidad de la base de datos	37
Tabla N° 8: Tabulación de conocimiento de metodologías de migración.....	38
Tabla N° 9: Tamaño de la base de datos	39
Tabla N° 10: Tabulación problemas con la información migrada	40
Tabla N° 11: Tabulación existencia de documentación de base de datos.	41
Tabla N° 12: Tabulación respecto a análisis de información realizados.	42
Tabla N° 13: Tabulación se realiza procedimientos de depuración en la base de datos....	43
Tabla N° 14: Tabulación tipos de aplicaciones que interactúan con la base de datos	44
Tabla N° 15: Tabulación datos crecimiento de la base de datos	45
Tabla N° 16: Tabulación existencia de inconsistencias en procesos migratorios previos.	46
Tabla N° 17: Frecuencias observadas.....	48
Tabla N° 18 : Frecuencias esperadas.....	49
Tabla N° 19: Cálculo de Chi-cuadrado	49
Tabla N° 20: Gastos empleados en recurso humano.	56
Tabla N° 21: Gastos Adicionales	56
Tabla N° 22: Características de metodología Strategy 7 Corporation	70
Tabla N° 23: Características metodología migración Oracle-sqlServer-SSMA(1)	71
Tabla N° 24: Características metodología migración Oracle-sqlServer-SSMA(1)	72
Tabla N° 25: Características de metodología de migración Álvaro Herrera (1)	73
Tabla N° 26: Características de metodología de migración Álvaro Herrera (2)	74
Tabla N° 27: Características de metodología de migración Álvaro Herrera (3)	75
Tabla N° 28: Características de la metodología de migración de datos de Hudicka.....	77
Tabla N° 29: Resumen de fases de metodologías de migración	79
Tabla N° 30: Cuadro comparativo de metodologías de migración	80
Tabla N° 31: Cuadro de planificación.	88
Tabla N° 32: Ficha recolección información de tablas	90

Tabla N° 33: Ficha recolección de información de funciones	90
Tabla N° 34: Ficha recolección de información de procedimientos	91
Tabla N° 35: Ficha depuración de objetos de base de datos.	94
Tabla N° 36: Cuadro planificación de pruebas de carga.....	97
Tabla N° 37: Planificación de la migración (1)	105
Tabla N° 38: Planificación de la migración (2)	106
Tabla N° 39: Características de los gestores de base de datos (1)	108
Tabla N° 40: Características de los gestores de base de datos (2)	109
Tabla N° 41: Características de los gestores de base de datos (3)	110
Tabla N° 42: Características de los gestores de base de datos (4)	111
Tabla N° 43: Características de los gestores de base de datos (5)	112
Tabla N° 44: Herramientas de migración (1)	114
Tabla N° 45: Herramientas de migración (2)	115
Tabla N° 46: Herramientas de migración (3)	116
Tabla N° 47: archivo de configuración de Ora2PG.....	118
Tabla N° 48: Planificación de pruebas.	123

ÍNDICE DE GRÁFICOS

Gráfico N° 1: Relación Causa-efecto	4
Gráfico N° 2: Red de inclusiones conceptuales.....	14
Gráfico N° 3: Subcategorías de la Variable Independiente	15
Gráfico N° 4: Subcategorías de la variable dependiente	16
Gráfico N° 5: Proceso de información	17
Gráfico N° 6: Niveles de abstracción de base de datos	19
Gráfico N° 7: Gestor de base de datos.....	22
Gráfico N° 8: Tabulación origen del proceso de migración	36
Gráfico N° 9: Tabulación complejidad de la base de datos	37
Gráfico N° 10: Tabulación Conocimiento de metodologías de migración	38
Gráfico N° 11: Tamaño de la base de datos	39
Gráfico N° 12: Problemas con la información	40
Gráfico N° 13: Existencia de documentación de base de datos	41
Gráfico N° 14: Tabulación se realiza análisis de información.....	42
Gráfico N° 15: Se realiza procedimientos de depuración de base de datos.	43
Gráfico N° 16: tipos de aplicaciones que interactúan con la base de datos	44
Gráfico N° 17: Crecimiento anual de base de datos.	45
Gráfico N° 18: Existencia de inconsistencias en procesos migratorios previos.	46
Gráfico N° 19: Verificación de hipótesis	50
Gráfico N° 20: Estructura de la Base de datos Oracle	57
Gráfico N° 21: Instancia Oracle	58
Gráfico N° 22: Principales componentes de postgresQL	63
Gráfico N° 23: Esquema de conexión postgresQL	65
Gráfico N° 24: Tipos de Migración en EnterpriseDB (Online y offline).....	67
Gráfico N° 25: Metodología de migración	85
Gráfico N° 26: Proceso de preparación de entorno de migración.....	87
Gráfico N° 27: Proceso de ejecución de la migración	95
Gráfico N° 28: Proceso de evaluación y control de calidad	99
Gráfico N° 29: Modelo Operativo	104
Gráfico N° 30: Ejecución de Ora2Pg	119
Gráfico N° 31: Verificación de creación de script.....	119
Gráfico N° 32: Extracción de triggers	120

Gráfico N° 33: Extracción de funciones	120
Gráfico N° 34: Extracción de procedimientos	120
Gráfico N° 35: Archivos creados por ora2Pg	121
Gráfico N° 36: Cliente de PostgreSQL pgAdminIII	121
Gráfico N° 37: Verificación de creación de tablas	122
Gráfico N° 38: Administración	125
Gráfico N° 39: Pantalla inicial de instalación herramienta SQLWays	147
Gráfico N° 40: Pantalla de configuración del origen en SQLWays	148
Gráfico N° 41: Pantalla de selección del destino de migración de SQLWays	148
Gráfico N° 42: Pantalla de selección de objetos de migración de SQLWays	149
Gráfico N° 43: Pantalla de instalación de SQLWays	150
Gráfico N° 44: Pantalla de instalación de Postgres Plus	151
Gráfico N° 45: Pantalla de instalación de componentes EnterpriseDB	151
Gráfico N° 46: Pantalla de instalación de configuración puerto Enterprise DB	152
Gráfico N° 47: Variables de entorno del sistema	153
Gráfico N° 48: Descarga de módulo DBD:Pg	154
Gráfico N° 49: Compilación módulo DBD::Pg	155
Gráfico N° 50: instalación módulo DBD::Pg	155
Gráfico N° 51: Descarga del módulo DBD::Oracle	155
Gráfico N° 52: Compilación módulo DBD::Oracle	156
Gráfico N° 53: creación de ejecutable	156
Gráfico N° 54: Ejecución de Test de módulo Oracle	157
Gráfico N° 55: Instalación de módulo Oracle	157
Gráfico N° 56: Compilación de Ora2Pg	158
Gráfico N° 57: Instalación de Ora2Pg	158
Gráfico N° 58: Instalación de Ora2Pg	159
Gráfico N° 59: Ejecución de scripts en PostgreSQL	160

UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA E
INDUSTRIAL
MAESTRIA EN GESTION DE BASES DE DATOS

“METODOLOGÍA PARA MIGRACIÓN DE DATOS QUE PERMITA
ASEGURAR Y CONSERVAR LA INTEGRIDAD Y CONSISTENCIA DE LA
INFORMACIÓN ADMINISTRADA POR LA EMPRESA VSYSTEMS”

Autor: Ing. Susana Beatriz Caraguay Martínez

Tutor: Ing. Mg. Franklin Oswaldo Mayorga Mayorga

Fecha: 17 de Septiembre de 2012

RESUMEN

Buscando un ahorro por costos de licenciamiento y la posibilidad de incursionar en el uso de software libre la empresa VSYSTEMS de la provincia de Santo Domingo de los Tsáchilas, decidió emprender un proceso de migración a nivel de aplicaciones y base de datos.

El problema principal de ésta empresa en su proceso de migración es que carece de una metodología de migración de base de datos que asegure que la información administrada conserve la integridad y exactitud luego de haber intervenido en un proceso de migración, así como la funcionalidad y rendimiento adecuado del nuevo gestor.

La presente tesis propone establecer una metodología de migración de datos que asegure la conservación de la integridad de la información, realizando el proceso de migración de la base de datos que interactúa con sistema de punto de venta e inventario desarrollado por la empresa VSYSTEMS.

Descriptores:

Migración de datos, metodología de migración de base de datos, integridad, gestores de base de datos.

TECHNICAL UNIVERSITY OF AMBATO

**FACULTY OF ENGINEERING SYSTEMS, ELECTRONICS AND
INDUSTRIAL**

MASTERS DEGREE IN DATABASE MANAGEMENT

“DATA MIGRATION METHODOLOGY TO ENSURE AND MAINTAIN THE INTEGRITY AND CONSISTENCY OF INFORMATION MANAGED BY THE ENTERPRISE VSYSTEMS”

Author: Engineer Susana Beatriz Caraguay Martínez

Tutor: Engineer. Mg. Franklin Oswaldo Mayorga Mayorga

Date: September 17, 2012

SUMMARY

To save on licensing cost and promote the use of free software, VSYSTEMS company located in Santo Domingo de los Tsáchilas, decided to undertake a process of migration at the application and database.

The main problem with this company in the process of migration is that it lacks a methodology for database migration to ensure that the information given conserve the integrity and accuracy after being involved in a migration process, as well as the functionality and adequate performance of the new DataBase Management System.

To solve the problem described, this thesis proposes to establish a data migration methodology that ensures the preservation of the integrity of the information, making the migration process of the database system point of sale and inventory of the company VSYSTEMS.

Descriptors:

Data migration, Database migration methodology, database management system

INTRODUCCIÓN

El trabajo de investigación tiene como tema: “METODOLOGÍA PARA MIGRACIÓN DE DATOS QUE PERMITA ASEGURAR Y CONSERVAR LA INTEGRIDAD Y CONSISTENCIA DE LA INFORMACIÓN ADMINISTRADA POR LA EMPRESA VSYSTEMS”. La importancia de ésta investigación radica en que proporcionará una metodología que asegure que la información que se migre entre gestores distintos, brinde como resultado una base de datos que conserve la integridad, permitiendo el funcionamiento correcto de los sistemas que hacen uso de esta información y evitar así información engañosa o incorrecta que repercute en graves pérdidas económicas para las empresas que necesiten realizar una migración de gestores de base de datos.

El Plan de Tesis está estructurado por capítulos de la siguiente manera:

El Capítulo I: EL PROBLEMA hace referencia al contexto y la problemática existente en la ejecución de procesos de migración entre gestores de base de datos en la empresa VSYSTEMS. Se realiza un análisis crítico y la formulación del problema que permite al final del capítulo establecer los objetivos del presente plan de tesis.

El Capítulo II llamado MARCO TEÓRICO contiene las bases teóricas, científicas, filosóficas, tecnológicas y legales que fundamentan la realización del presente plan de tesis. Además se presenta el planteamiento de hipótesis.

En el Capítulo III llamado METODOLOGÍA se describe la información que hace referencia al proceso de investigación realizado y las técnicas empleadas en dicho proceso. Contiene también la planificación de recolección de información y plan de procesamiento de información.

El Capítulo IV ANÁLISIS E INTERPRETACIÓN DE RESULTADOS presenta los resultados obtenidos en la recolección y tabulación de la información que permite realizar la verificación de hipótesis del presente plan de tesis. Consta de los siguientes contenidos: Análisis de los resultados, interpretación de datos y verificación de hipótesis.

El Capítulo V CONCLUSIONES Y RECOMENDACIONES, plasma las conclusiones y recomendaciones luego de haber realizado el análisis e interpretación de resultados. Estas conclusiones y recomendaciones están orientadas a resolver el problema planteado en el presente plan de tesis.

En el Capítulo VI PROPUESTA, se plantea una alternativa de solución a la problemática de migración de datos entre diferentes gestores en la empresa VSYSTEMS. Se describe un modelo operativo y se plantea una metodología de migración de datos para la empresa VSYSTEMS. Está formado por los siguientes temas: antecedentes, justificación, objetivos, análisis de factibilidad, fundamentación, metodología, modelo operativo, administración y previsión de la evaluación.

CAPÍTULO I

1. EL PROBLEMA

1.1. Tema

La aplicación de una metodología establecida para la migración de datos entre gestores de base de datos y su influencia en la conservación de la integridad y la consistencia de información.

1.2. Planteamiento del problema

1.2.1. Contextualización

El principal activo de las empresas es la información resultante de sus actividades propias, que al ser administrada adecuadamente permite toma de decisiones y facilita el alcance de los objetivos empresariales. Ésta información se administra mediante sistemas informatizados y es almacenada en gestores de base de datos.

La migración de datos surgió con la necesidad de que la información pueda ser portable entre diferentes gestores de Bases de Datos. Uno de los principales problemas que pueden ocurrir al momento de realizar un proceso de este tipo sin la preparación y ejecución adecuada es que se ocasione pérdida de información, que conduciría a que las empresas tengan graves problemas en el desempeño de sus actividades al no contar con la información correcta y precisa en sus sistemas repercutiendo en graves pérdidas económicas.

Otro factor que ha incrementado los procesos de migración de Base de datos está relacionado con la tendencia actual de los países latinoamericanos hacia el uso de software libre a nivel de gestores de Base de datos como a nivel de

aplicación de usuario final, basando dicho cambio a la real ventaja del Software Libre que la adquisición del programa permite libertad al usuario, desde la parte funcional permite que se instale en el número de máquinas que se requiera, sin la necesidad de pagar por nuevas licencias de uso. Entre las Base de datos de software libre más usadas se encuentran PostgreSQL y FireBird.

En nuestro país la tendencia hacia la innovación tecnológica junto al crecimiento empresarial, conlleva a las empresas a buscar mejoras sus sistemas informáticos y procesos de administración de información incluyendo en ello a los gestores de base de datos, y en otros casos con el objetivo de reducir costos por pagos de licencia se busca alternativas de gestores que cubran los requerimientos empresariales dando así inicio a una necesidad de migración de información. Además en nuestro país se ha determinado a nivel gubernamental decretos que aseguran el uso de software libre. En el Ecuador el Gobierno ha promulgado el uso del software libre, a través del Decreto 1014, emitido el 10 de Abril del 2008 por la Presidencia de la República. Un extracto del decreto contiene los siguientes artículos:

Art. 1: Establecer como política pública para las entidades de administración Pública central la utilización del Software Libre en sus sistemas y equipamientos informáticos.

Art. 2: Se entiende por software libre, a los programas de computación que se pueden utilizar y distribuir sin restricción alguna, que permitan el acceso a los códigos fuentes y que sus aplicaciones puedan ser mejoradas.

Estos programas de computación tienen las siguientes libertades:

- *Utilización de programa con cualquier propósito de uso común.*
- *Distribución de copias sin restricción alguna*
- *Estudio y modificación de programa (Requisito: código fuente disponible)*
- *Publicación del programa mejorado (Requisito: código fuente disponible)*

Art. 3: Las entidades de la administración pública central previa a la instalación del software libre en sus equipos, deberán verificar la existencia de capacidad técnica que brinde el soporte necesario para este tipo de software.

Art. 4: Se faculta la utilización de software propietario (no libre) únicamente cuando no exista una solución de software libre que supla las necesidades requeridas, o cuando esté en riesgo de seguridad nacional, o cuando el proyecto informático se encuentre en un punto de no retorno.

1.2.2. Árbol de problemas

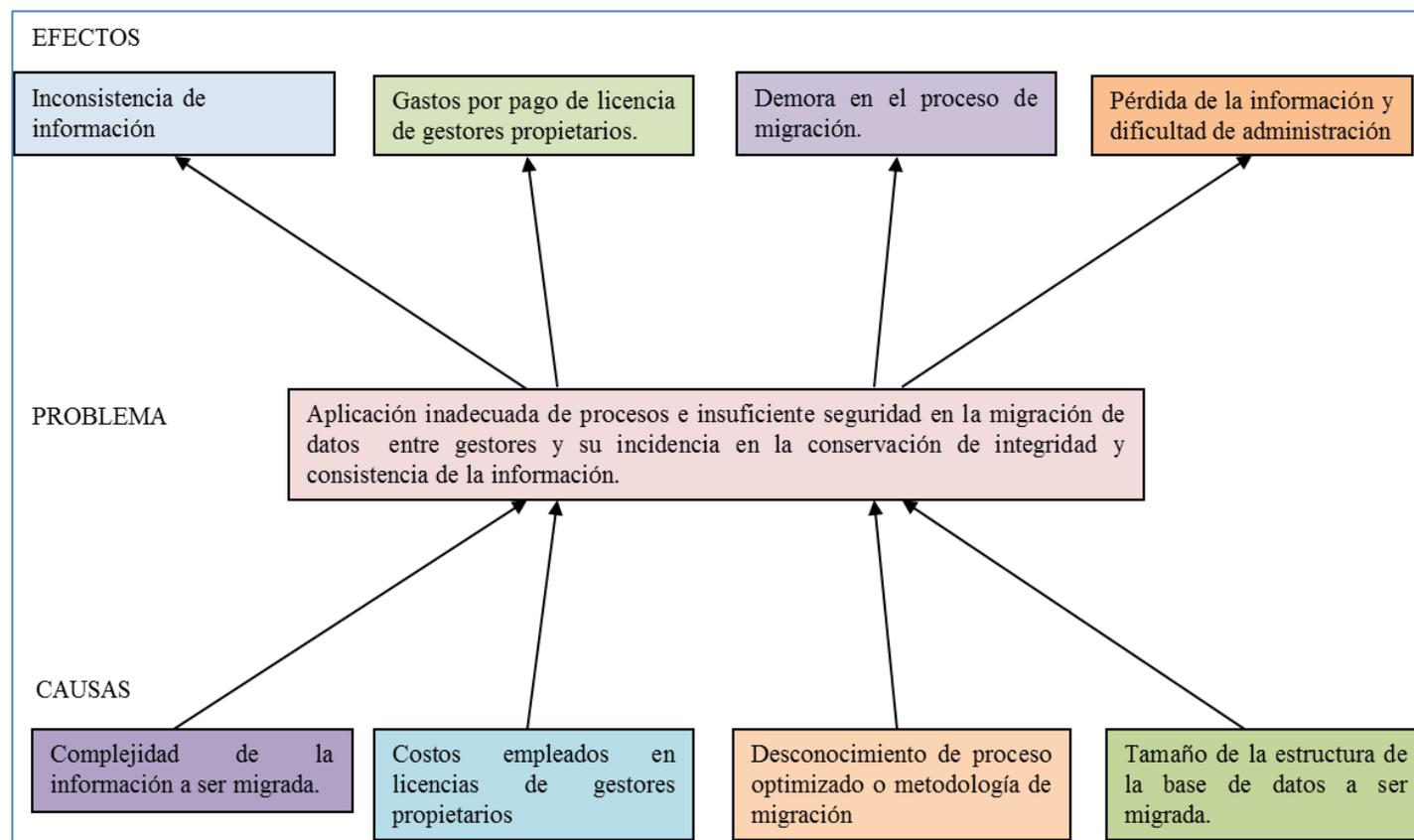


Gráfico N° 1: Relación Causa-efecto

Fuente: Investigadora

1.2.3. Análisis Crítico

La información constituye actualmente uno de los recursos más importantes que poseen las empresas ya que guían a la toma de decisiones y permiten gestionar de manera óptima y fácil las actividades empresariales.

Para la administración de la información se utilizan los gestores de base de datos encargados de organizar, almacenar, proveer rutinas para manejar transacciones correspondientes a información relevante y vital para el funcionamiento de las aplicaciones de las empresas.

Existen diferentes gestores de base de datos y son las empresas quienes seleccionan el gestor que se adapte a sus necesidades, dando como resultado muchas veces el cambio del gestor actual por uno con nuevas características y funcionalidades, generando así un proceso de migración de datos.

Entre los factores considerados para decidir un cambio de gestor de base de datos se encuentran:

- Reducir costos por licenciamiento.
- Cambiar a alternativas de base de datos de código abierto
- Cambiar a gestores que tienen mejoras continuas y cuentan con comunidades web de usuarios que dan soporte.
- Crecimiento de la base de datos y el gestor actual no rinde el beneficio esperado para manejar la cantidad de información almacenada.
- Incremento en complejidad de procesos.

Los costos que son empleados anualmente en la adquisición de licencias de gestores de Base de datos propietario producen incremento en el gasto de las instituciones, por lo cual la mejor opción para reducir los costos es realizar una migración hacia un gestor de software libre que permite instalación ilimitada,

brindando al mismo tiempo acceso a código abierto, permitiendo ahorrar recursos económicos a la empresa.

El desconocimiento de un proceso óptimo de migración ocasiona demora en los procesos. La falta de planificación y de calidad durante un proceso de migración puede volver a la base de datos inconsistente, ya que al no tener un proceso correctamente definido se implementa soluciones imprevistas y poco ortodoxas que necesitarán posterior revisión minuciosa para determinar el nivel de consistencia teniendo en muchos de los casos que volver a realizar todo el proceso causando grandes retrasos.

La administración de la información demanda que se establezcan dentro de las bases de datos un conjunto de relaciones entre sus datos, restricciones entre los mismos, funciones, procedimientos y triggers que aseguran la integridad de la información. Si el momento de realizar una migración no se asegura que dichas relaciones y restricciones se conserven, tendría un efecto negativo pues la información que se almacenaría podría no ser la real y causaría dificultades a las aplicaciones que administran dicha información como a las empresas propietarias de la información.

El tamaño de la base de datos a migrar también es un factor determinante en un proceso de migración, al aumentar el tamaño es mayor la dificultad en el proceso de migración y posterior validación de información.

1.2.4. Prognosis

El crecimiento continuo de información y la complejidad para su administración da como resultado la búsqueda y cambio de gestores de base de datos, eligiendo opciones que se adapten a las necesidades de las empresas dando origen a los procesos de migración de datos.

Si no se establece una metodología adecuada para la migración ante la complejidad de la información a ser migrada, se puede llegar a una incoherencia total entre la información real y la almacenada en la base de datos dificultando el acceso de las aplicaciones empresariales.

En caso de no aplicar un proceso óptimo de migración se produciría un alto desperdicio de recurso humano y tiempo, representando pérdidas económicas para la empresa.

De continuar con la adquisición de licencias de gestores de software propietario se produce un incremento al gasto anual perjudicando las finanzas de la institución, el presupuesto que se invierte en el pago de licencias puede ser alto para el presupuesto que maneja la empresa.

1.2.5. Formulación del problema

¿Cómo incide en la conservación de integridad y consistencia de la información la aplicación de una metodología específica y uso de una herramienta de migración de datos entre gestores Oracle10g a PostgreSQL en la empresa VSYSTEMS?

1.2.6. Interrogantes de la investigación

- ¿Qué función cumple la migración de datos?
- ¿Cuáles son los subprocesos que intervienen en la migración de datos?
- ¿Qué técnicas son los más recomendados para realizar la migración de datos?
- ¿Qué características debe poseer una herramienta de migración para asegurar el un proceso óptimo?
- ¿Cómo influye la metodología de migración empleada en la conservación de integridad y consistencia de la información?

1.2.7. Delimitación del objeto de investigación.

1.2.7.1. Delimitación de Contenido

Campo: Sistemas

Área: Gestión de Base de Datos

Aspecto: Migración de datos y Seguridad

1.2.7.2. Delimitación espacial

La investigación se desarrolló en la empresa VSYSTEMS ubicada en la ciudad de Santo Domingo, Provincia de Santo Domingo de los Tsáchilas.

1.2.7.3. Delimitación Temporal

El trabajo de investigación se realizó durante el periodo Julio 2011- Julio 2012.

1.2.7.4. Unidad de observación

- Usuarios de aplicaciones que se conectan a la Base de datos
- Usuarios avanzados de base de datos.
- Administrador de Base de datos.

1.3. Justificación

El trabajo de investigación científica tiene **utilidad teórica** porque contribuye con nuevos conocimientos en base a información actualizada y especializada sobre el proceso de migración de datos al definir un proceso óptimo para realizar la migración.

La investigación tiene Utilidad Práctica por qué se plantea una alternativa de solución al problema a investigar. En ésta investigación se obtiene una

metodología que facilita y garantiza la aplicación efectiva del proceso de migración.

La presente investigación es **factible** ya que elabora la propuesta de un modelo operativo para solucionar el problema planteado, mediante el uso de recursos tecnológicos y bibliográficos actualizados y especializados, así como el conocimiento de la investigadora sobre el tema propuesto.

1.4. Objetivos

Objetivo General

Establecer una metodología de migración de datos que asegure la conservación de integridad y consistencia de información administrada por la empresa VSYSTEMS.

Objetivos Específicos

- Analizar y seleccionar una herramienta de migración de datos desde un gestor Oracle 10g a PostgreSQL que brinde características adecuadas para una migración óptima.
- Plantear una alternativa de solución al problema de insuficiente seguridad en el proceso de migración de datos desde gestor Oracle10g a PostgreSQL y su incidencia en la conservación de integridad y consistencia de la información.
- Definir una metodología para el proceso de migración aplicable a un modelo entidad relación analizando los factores que determinan la consistencia e integridad de información.
- Aplicar la metodología de migración y herramienta en un caso de prueba específico en un modelo entidad relación de la empresa VSYSTEMS.

CAPÍTULO II

2. MARCO TEÓRICO

2.1. Antecedentes Investigativos

Actualmente en algunos países de Latinoamérica se están realizando estudios referentes al tema de migración de datos y aplicaciones. En Chile la Universidad Tarapacá de Arica realizó un estudio de factibilidad de Migración a Software Libre de todas sus aplicaciones y Bases de datos, los resultados obtenidos concluyen que la migración a Software libre es factible obteniendo la organización independencia tecnológica es decir libertad para desarrollar y administrar sus sistemas computacionales de la forma en que estime conveniente según sus propios intereses y no según los intereses de las empresas de software privada.

En Venezuela el año 2005 el Centro Nacional de Tecnologías de la Información perteneciente al ministerio del poder popular para las telecomunicaciones y la informática lanza el plan nacional de migración a software libre de la administración pública nacional, cuyo objetivo es emplear los mecanismos necesarios para aplicar estándares abiertos en el uso y desarrollo de software libre en la Administración Pública Nacional. Involucrando por lo tanto el cambio de gestores de base de datos propietarios a libres.

En el Ecuador la Subsecretaria de informática en el año 2009, desarrolla la estrategia para la implantación de software libre en la administración pública central, en Quito 2009, incluyendo dentro del punto 4.3.3 la migración de base de datos al especificar un orden de migración como se detalla a continuación:

- Servidor de aplicaciones

- Servidor de correo electrónico
- Servidor de archivos e impresión
- Servidor Web, en el caso de páginas estáticas.
- Servidor Web y base de datos en el caso de páginas dinámicas.
- Servidor de seguridades
- Directorio activo
- Servidor de base de datos

Adicionalmente la Subsecretaria de Informática del Ecuador en el año 2009 publicó en su página web un listado de alternativas de software libre a las que se puede migrar las aplicaciones y bases de datos desde software privado a software libre. Entre las bases de datos y sus alternativas se encuentran:

Base de datos propietarias	Base de datos de software libre
Microsoft SQL, Oracle	Postgres, MySQL

Tabla N° 1: Alternativa de software libre para base de datos
Fuente: Subsecretaria de Informática del Ecuador (2009)

2.2. Fundamentaciones

2.2.1. Fundamentación Filosófica

Para realizar el trabajo de grado la investigadora asume los principios del paradigma crítico-propositivo, por que cuestiona y analiza aspectos basados en la existencia de múltiples realidades socialmente construidas y plantea una propuesta de solución al problema investigado

2.2.2. Fundamentación Tecnológica

El trabajo de investigación tiene su fundamento en la tecnología por que permite la optimización de procesos de migración y asegura la consistencia de la

información., mediante el uso de herramientas y técnicas informáticas para llegar a la solución del problema.

2.2.3. Fundamentación Legal

La investigación tendrá su base legal en el Art. No. 37 del **REGLAMENTO CODIFICADO DE RÉGIMEN ACADÉMICO DEL SISTEMA NACIONAL DE EDUCACIÓN SUPERIOR del 22 de enero de 2009, el mismo que dice:** “Art. 37.5 Para obtener el grado de Magíster, los postulantes deben realizar y sustentar una tesis de investigación científica que presente novedad y originalidad en el problema, los materiales de investigación, los métodos aplicados y en las conclusiones y recomendaciones”.

Otra fundamentación legal se encuentra en la Resolución adoptada por el CONESUP en la sesión ordinaria No.17 del día 26 de Junio del 2008 mediante la resolución denominada RCP.S17.No.191.08 que dice: “Consejo Nacional de Educación Superior, luego del análisis y debate correspondiente resuelve aprobar el proyecto de postgrado “Maestría en Gestión de Bases de Datos”, presentado por la Universidad Técnica de Ambato, para desarrollarse en la ciudad de Ambato, sede de la Universidad, en modalidad de estudios semipresencial.”

2.3. Categorías Fundamentales

2.3.1. Red de inclusión

La fundamentación teórica de la presente investigación se enfoca en los aspectos que señala el gráfico N° 2.

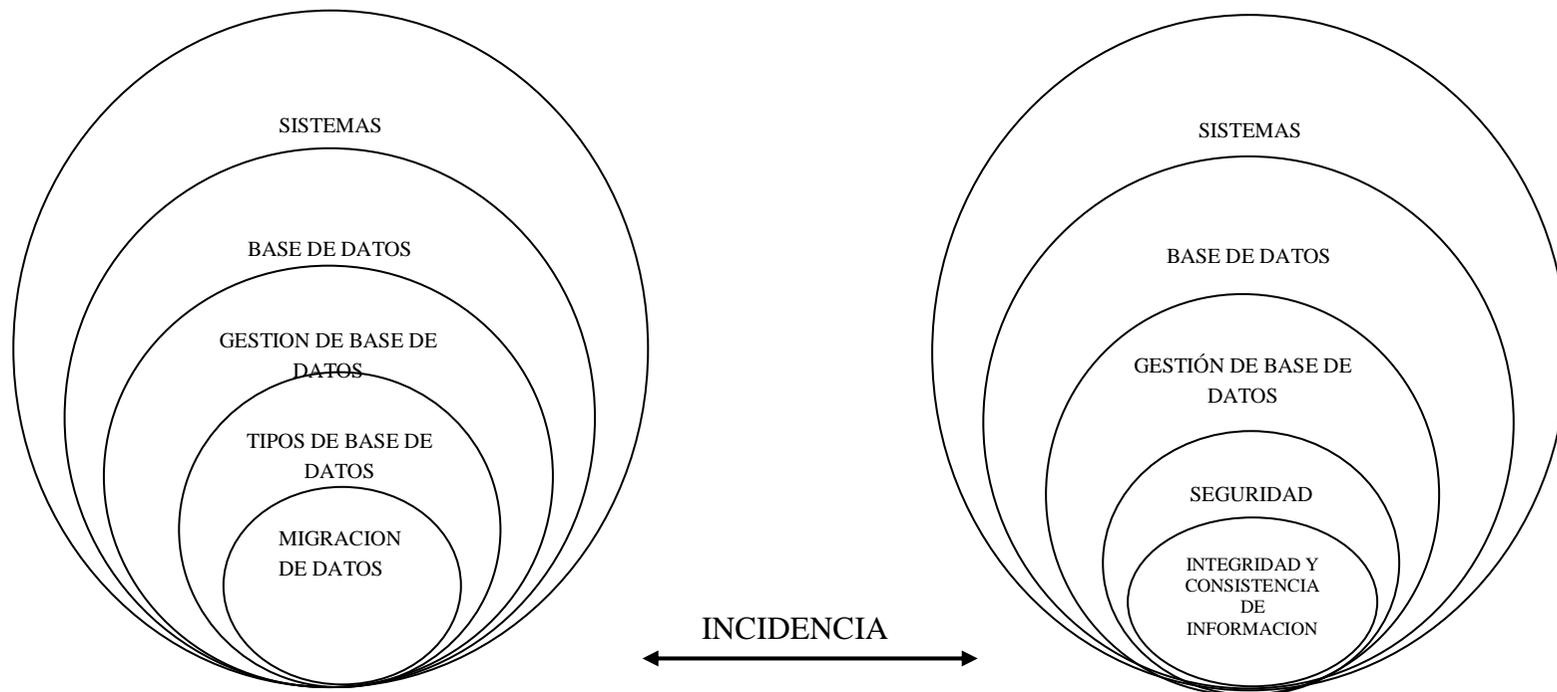


Gráfico N° 2: Red de inclusiones conceptuales.

Fuente: Investigadora

2.3.2. Constelación de ideas de la Variable Independiente

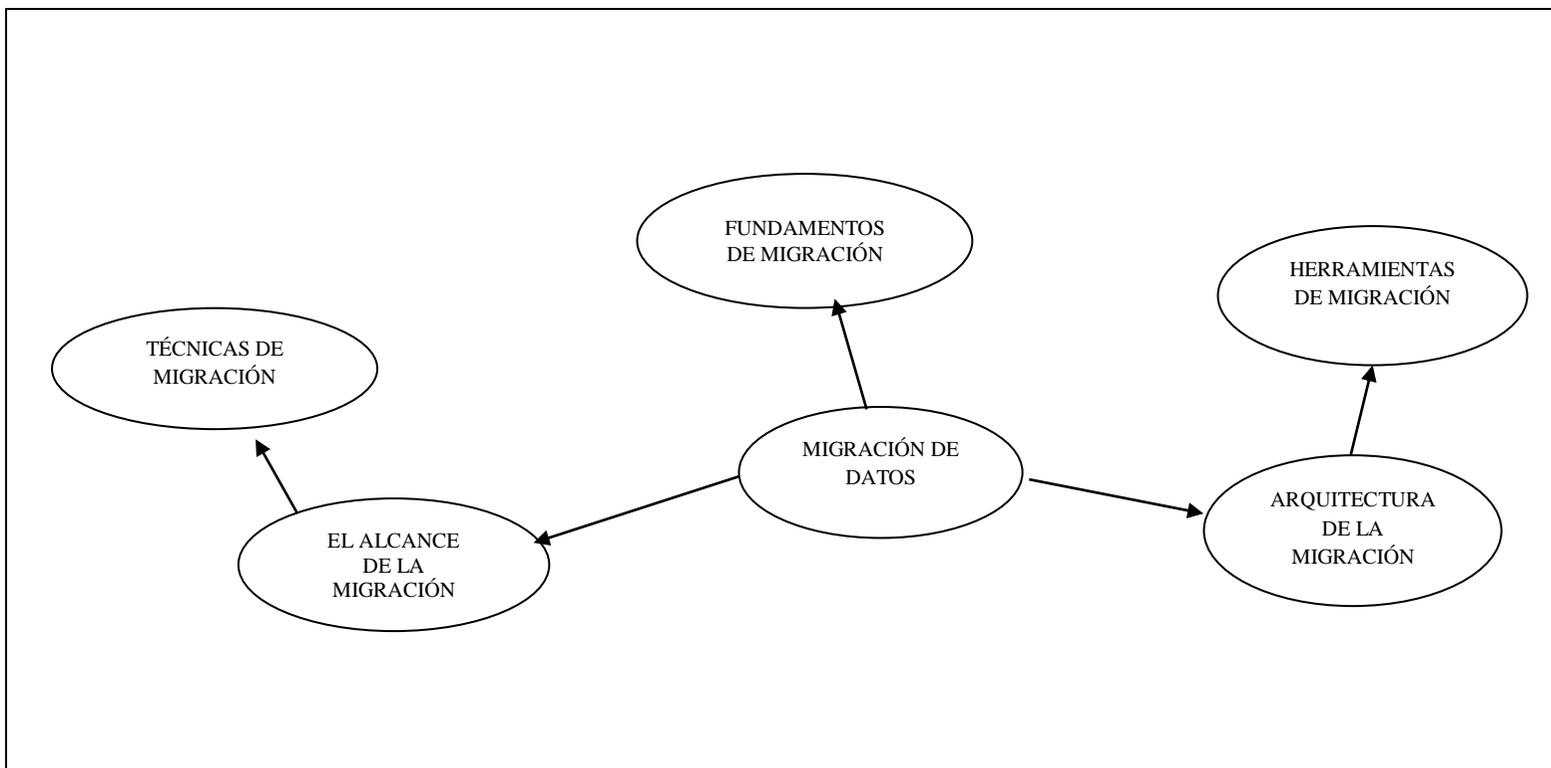


Gráfico N° 3: Subcategorías de la Variable Independiente

Fuente: Investigadora

2.3.3. Constelación de ideas de la Variable Dependiente

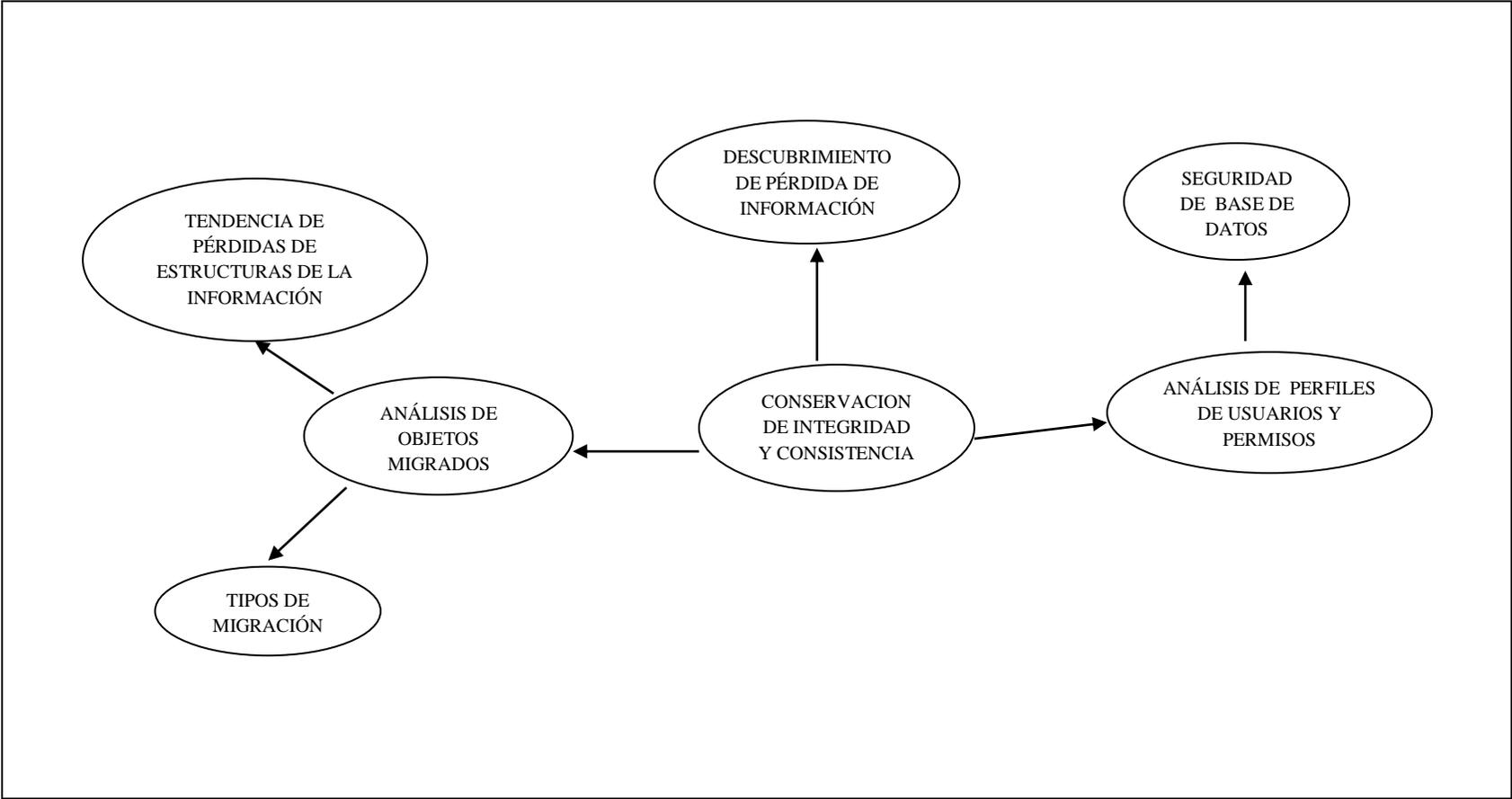


Gráfico N° 4: Subcategorías de la variable dependiente

Fuente: Investigadora

SISTEMAS DE INFORMACIÓN

Un sistema de manera general se puede definir como un conjunto de entidades cuya interacción es orientada a lograr un objetivo.

Se puede describir a un sistema de información como la agrupación de elementos interrelacionados para administrar, procesar o recuperar la información que es obtenida a partir de datos.

Como todo sistema, los sistemas de información poseen estructura y/o funcionalidad, se encuentran en un entorno específico y tienen límites. El gráfico N° 5 representa el proceso que da origen a la información.

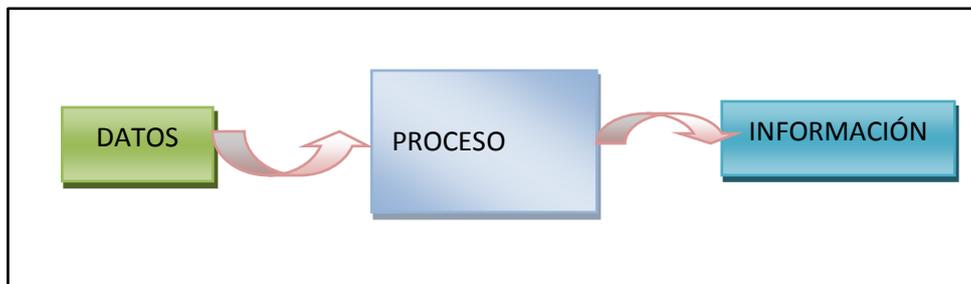


Gráfico N° 5: Proceso de información

Fuente: Investigadora

Actualmente toda organización posee sistemas de información para el desarrollo de sus actividades, constituyéndose en uno de los activos más preciados, por lo que su correcta administración es un factor clave en el desempeño de las tareas organizacionales.

Con la finalidad de contar con un almacenamiento permanente y organizado para los sistemas de información, fueron creados diferentes dispositivos y métodos de almacenamiento.

La interacción entre las nuevas formas de almacenamiento y técnicas de administración de la información dió como resultado el origen de las bases de datos.

BASE DE DATOS

Una base de datos es un conjunto de información ordenada y almacenada en estructuras específicas para satisfacer las necesidades de información de las organizaciones. La mayor parte de bases de datos son almacenadas en plataformas informáticas para su administración y gestionadas a través de aplicaciones y rutinas.

Niveles de abstracción de las bases de datos

Con el fin de proporcionar a los usuarios de la base de datos una visión abstracta de la información, ocultando detalles de almacenamiento y permanencia, se establecen diferentes niveles de abstracción en una base de datos.

Se considera los siguientes niveles de abstracción de base de datos:

- Nivel Externo: Es el nivel de abstracción que describe como es vista la información por los usuarios finales.
- Nivel Conceptual: Es la representación de los datos y sus relaciones mediante estructuras conceptuales que representan la realidad que se intenta replicar a la base de datos
- Nivel Físico: El nivel de abstracción físico hace referencia a la forma como se almacenará físicamente la información así como rutas de acceso y compresión de datos.

El gráfico N° 6 representa los diferentes niveles de abstracción de la base de datos.

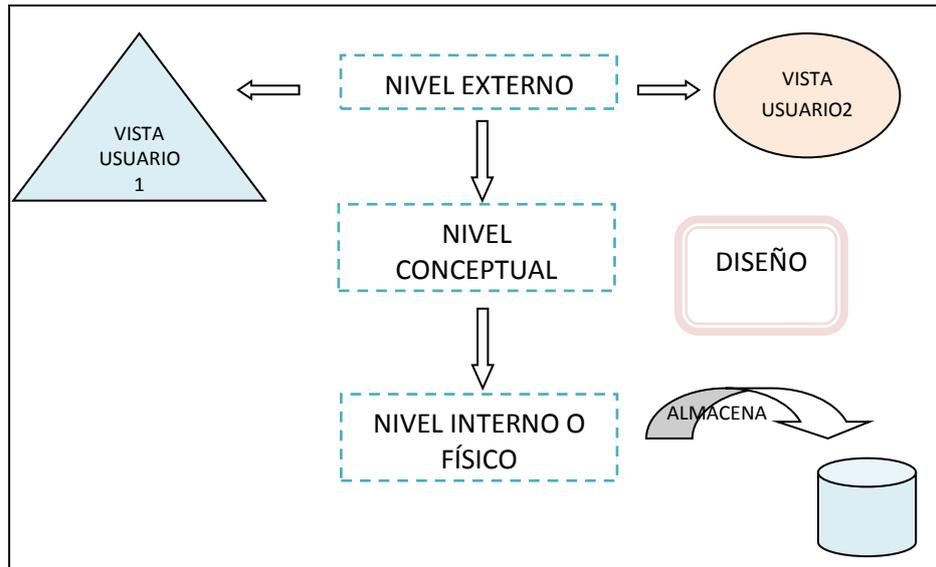


Gráfico N° 6: Niveles de abstracción de base de datos

Fuente: Investigadora

Modelos de datos

Un modelo de datos es una forma de representar y explicar de manera lógica la organización y relaciones de los datos. Existen algunos modelos de datos entre los principales se encuentran los siguientes:

- Base de datos jerárquica su representación es construida con una raíz y varios subniveles a manera de árbol.
- Base de datos relacionales, Codd (1970) describió el modelo relacional en el cual los datos y asociaciones entre ellos son representados como relaciones (tablas)
- Base de datos entidad/relación es la base de datos en la cual la información es organizada como entidades que representan la realidad, existiendo vinculaciones o interrelaciones entre las entidades
- Base de datos orientada a objetos, la información es representada mediante objetos, es decir aplicando el paradigma de orientación a objetos con características de encapsulación, herencia y polimorfismo.

GESTIÓN DE BASE DE DATOS

La gestión de base de datos consiste en la realización de actividades por parte del administrador de base de datos para garantizar el correcto almacenamiento, uso, validez y permanencia de la información. Para ello se hace uso de tecnologías de almacenamiento, sistemas gestores de base de datos, protocolos de red e infraestructura tecnológica. Entre las principales actividades que se realiza dentro de la gestión de base de datos se encuentran:

- Diseño y creación de base de datos
- Configuración de base de datos
- Administración de usuarios
- Control de acceso
- Soluciones de almacenamiento
- Generación de backups y recuperación
- Optimización de base de datos
- Seguridad, auditoria y monitoreo de actividades
- Soluciones de minería de datos

En las bases de datos administradas en sistemas de cómputo, el manejo de información es realizado por el sistema Gestor de base de datos (SGBD).

El SGBD está conformado por datos interrelacionados y el conjunto de programas que permite tener acceso a dichos datos usando un lenguaje propio de base de datos.

El sistema gestor de base de datos sirve de interfaz para la administración de los datos y los objetos de la base de datos.

Estructura de un sistema gestor de Base de datos

La estructura de un sistema gestor de base de datos hace referencia a los componentes que forman parte de ella y permiten su funcionamiento.

Basándonos en la información de Khort, Silberschartz & Sudarshan (2006) se identifica los siguientes elementos de la base de datos:

Gestor de almacenamiento

El gestor de almacenamiento es el responsable de la interacción de los datos con el sistema de archivos, cuenta con los siguientes módulos:

- Gestor de autorización e integridad: Gestiona acceso de usuarios a datos y la integridad de los datos.
- Gestor de transacciones: Asegura el estado consistente de una base de datos luego de realizar las transacciones.
- Gestor de archivos: Se encarga de la administración y reserva de espacio en disco para las estructuras de la base de datos.
- Gestor de memoria intermedia, traslada datos del disco a la memoria principal.

Para cumplir con su función el gestor de almacenamiento maneja estructuras de base de datos, tales como:

- Archivos de datos
- Diccionario de datos
- Índices

Procesador de consultas

El procesador de consulta permite al usuario interactuar con la información almacenada en la base de datos mediante sentencias escritas en lenguajes de consulta.

Las funciones que cumple el procesador de consultas de un sistema gestor de base de datos son las siguientes:

- Intérprete de LDD (Lenguaje de definición de datos), convierte las sentencias LDD y guarda las definiciones en el diccionario de datos

- Compilador de LMD (Lenguaje de manipulación de datos), realiza la evaluación de las sentencias y las traduce a un lenguaje de consulta.
- Motor de consultas es el que ejecuta las sentencia una vez que han sido compiladas

El gráfico N° 7 describe la funcionalidad del gestor de base de datos.

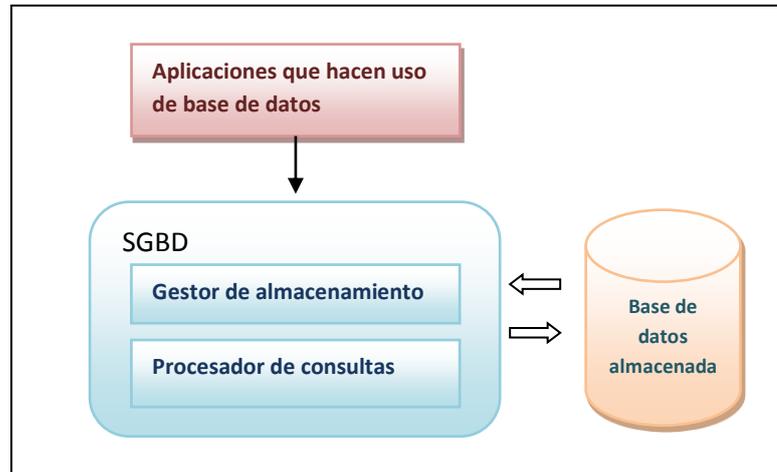


Gráfico N° 7: Gestor de base de datos

Fuente: Investigadora

SEGURIDAD DE BASE DE DATOS

La seguridad de la base de datos está relacionada a la protección de la información contra toda acción que ponga en riesgo la información almacenada así como su validez, algunos eventos perjudiciales para la información son:

- Robo de información crítica de la empresa
- Parálisis de operaciones
- Fraudes mediante alteración mal intencionada de la información.
- Alteración de la información produciendo invalidez.

Para determinar y controlar los riesgos existen algunas actividades que realizan los administradores de base de datos como:

Auditoría de base de datos: La auditoría de base de datos es el proceso mediante el cual se monitorea y analiza los accesos a la información y recursos dentro de la base de datos.

Criptografía: La criptografía es la técnica que se emplea para proteger la información mediante codificación, convirtiéndola en ilegible para usuarios no autorizados.

Control de acceso a recursos: El control de acceso a recursos es el proceso en el cual se asigna permisos específicos a los usuarios para que puedan hacer uso de la información y herramientas del gestor de base de datos. A cada usuario se le concede permisos dependiendo de su función o rol dentro de la organización. Solo usuarios con permiso para acceder a cierta información podrán hacerlo.

Objetivos de implantación de seguridad en la base de datos

Los objetivos de la implantación de seguridad en la base de datos están orientados hacia la protección de la información existente en ella.

Las características de seguridad indispensables para las bases de datos son:

- **Confidencialidad:** La información debe ser vista sólo por los usuarios que necesiten verla, para ello muchos gestores poseen funcionalidades para la autenticación de usuarios, autorización de usuarios y control de acceso a recursos.
- **Integridad:** Asegura la validez de la información almacenada en la base de datos, protege la información relacionada entre sí, implica aspectos como autorización de modificación de información a usuarios determinados.

- Disponibilidad: La información debe ser disponible para los usuarios en todo momento y circunstancia, por ello se debe analizar aspectos como la escalabilidad, flexibilidad de la base de datos y servidores de respaldo,

INTEGRIDAD Y CONSISTENCIA DE INFORMACION

De acuerdo con Silberschatz, Khort y Sudarshan (2002): *“Las restricciones de integridad proporcionan un medio de asegurar que las modificaciones hechas a la base de datos por los usuarios autorizados no provoquen la pérdida de la consistencia de datos.*

Para Miguel, Piattini y Marco (2000): *“La semántica y la integridad son conceptos de bases de datos en el cual se conoce como semántica al significado de los datos y con el de integridad a la corrección de los mismos y a su consistencia respecto al mundo real del cual proceden.”*

La integridad es la característica que asegura que la información sólo puede ser modificada por usuarios autorizados y con acceso al recurso asignado garantizando la completitud de la información y cumpliendo las reglas predefinidas respecto a las características de los datos.

Por otra parte la consistencia de la información tiene que ver con la correctitud de la información en un momento determinado, su coherencia y correspondencia con la realidad del modelo de datos que representa. Ésta correctitud tiene relación a su vez con el cumplimiento de integridad.

MIGRACIÓN DE BASE DE DATOS

La migración de base de datos puede definirse, como: el proceso mediante el cual se traslada la información entre gestores de base de datos.

La migración puede realizarse entre gestores homogéneos pero de diferente versión o entre gestores heterogéneos implicando en éste caso una transformación de tipos de datos y estructuras de la base de datos.

Para realizar un proceso de migración no existe una metodología específica estandarizada sino que de acuerdo a las necesidades las empresas se emplean técnicas de diferente tipo que muchas de las veces conllevan a un incremento en la duración del proceso de migración.

Los pasos que se podrían considerar como técnicas que deben formar parte de todo proceso de migración son:

- Planificación para realizar un proceso de migración se debe tener claro y especificado las actividades y tareas a realizar así como el plazo establecido para la realización de cada tarea.
- Mapeo de datos es la coincidencia entre estructuras de la base de datos.
- Extracción de datos desde la base de datos origen
- Carga de la información al nuevo gestor.

Metodología de migración de base de datos

De manera general se puede definir una metodología como el conjunto de procesos, tareas, técnicas y herramientas que se utilizan para realizar un proyecto específico.

Refiriéndonos exactamente a una metodología de migración de base de datos podemos describirla como el conjunto de procesos, ejecución de actividades, técnicas y uso de herramientas software para trasladar y/o convertir datos de un gestor de base de datos a otro.

2.4. Hipótesis

Al establecer una metodología para aplicar migración de datos se optimizará y asegurará la conservación de integridad y consistencia de la información administrada por la empresa VSYSTEMS.

Variable Independiente

Metodología para aplicar migración de datos

Variable Dependiente

Integridad y consistencia de información

CAPÍTULO III

3. METODOLOGÍA

3.1. Modalidades Básicas de Investigación

3.1.1. Investigación de campo

La investigación tiene la modalidad de campo porque la investigadora acudió al lugar en donde se produjeron los hechos para obtener información relacionada con los objetivos del trabajo de grado. Se realizó un análisis y procesos controlados con las variables para describir de qué modo o porque se produce una situación particular.

3.1.2. Investigación documental – bibliográfica

Con el propósito de fortalecer la investigación, se recurrió a obtener información teórica de diferentes autores obtenidas en fuentes secundarias (libros, revistas especializadas, publicaciones, internet, otros) y de fuentes de información primaria a través de documentos válidos y confiables.

3.1.3. Proyectos factibles de Intervención Social

El trabajo de grado responde a un proyecto factible de intervención social porque se plantea una propuesta viable de solución al problema investigado dentro de un contexto determinado que conlleve a una mejora de la realidad.

3.1.4. Proyectos Especiales

Con la finalidad de dar solución a los problemas de un contexto determinado la investigación tiene la modalidad de proyectos especiales porque con la utilización de la tecnología se construirá una metodología para aplicar migración de datos.

3.2. Tipos o Niveles de investigación

3.2.1. Exploratorio

La investigación pasó por el nivel de investigación exploratorio porque se realizó el sondeo de un problema poco investigado o desconocido en un contexto determinado, especialmente para definir el tema de investigación.

3.2.2. Descriptivo

Es descriptivo por que se busca informar los resultados obtenidos de la investigación entre la comparación de dos variables, tomando en cuenta criterios de coherencia interna y pertinencia.

3.2.3. Asociación de variables

Es de asociación de variables porque permite establecer y estudiar las tendencias de comportamiento entre variables en un contexto determinado.

3.3. Población y Muestra

La descripción de la población y muestra que fue considerada en la presente investigación se encuentra detallada en la tabla N° 2.

POBLACIONES	FRECUENCIA	PORCENTAJE
Administrador de Base de datos	1	16%
Usuarios avanzados de base de datos.	5	84%
TOTAL	6	100%

Tabla N° 2: Unidades de Observación

Fuente: Investigadora.

En virtud de que ninguna de las poblaciones a ser investigadas supera los cien elementos se trabajará con la totalidad del universo sin que sea necesario sacar muestras representativas.

3.4. Operacionalización de Variables

3.4.1. Operacionalización de Variable Independiente

Variable: Metodología para migración de datos

CONCEPTUALIZACIÓN	DIMENSIONES	INDICADORES	ITEMS BÁSICOS	TÉCNICAS E INSTRUMENTOS
Procedimiento mediante el cual se cambian los gestores actuales de Base de datos de una organización, por otros productos o versiones que se adapten mejor a las necesidades actuales y/o futuras de dicha organización de qué manera que la información pueda ser portable.	Procedimientos Innovación, necesidad de cambio Portabilidad de datos(Gestores de base de datos)	Diseño y documentación de procesos de base de datos. Rendimiento Agrupación o clustering. Disponibilidad Gestores propietarios, Gestores de Software Libre.	¿Cómo se puede determinar la técnica más conveniente ejecutar el proceso de migración exitoso? ¿Contribuye significativamente la migración de datos a preservar la información importante para la organización? ¿Cómo se determina la herramienta de migración que intervendrá en el proceso?	Técnica: Encuesta e Instrumento: Cuestionario.

Tabla N° 3: Operacionalización de la variable independiente.

Fuente: Investigadora

3.5. Técnicas e instrumentos

Encuesta: Dirigido al administrador de base de datos y operarios de la base de datos cuyo instrumento fue el cuestionario elaborado con preguntas cerradas para obtener información sobre el conocimiento de proceso de migración, complejidad de base de datos, tamaño de la base de datos.

Observación: Nos permite tener una percepción sistemática del proceso de manipulación de datos y la importancia de cada elemento en la migración.

Entrevista: Dirigido al administrador de base de datos Oracle 10G y PostgreSQL cuyo instrumento fue la guía de la entrevista, elaborada con preguntas abiertas para obtener información sobre procesos de migración. Entrevista al gerente de la empresa para obtener información sobre los objetivos empresariales tras el proceso de migración.

3.6. Validez y confiabilidad

Los instrumentos fueron sometidos a criterios de validez, a través de la técnica “juicio de expertos”, mientras que la confiabilidad fue verificada mediante la aplicación de una prueba piloto dirigida a una población pequeña para detectar errores y corregirlos a tiempo.

3.7. Plan de recolección de la información

PREGUNTAS BÁSICAS	EXPLICACIÓN
1. ¿Para qué?	Para alcanzar los objetivos de la investigación
2. ¿De qué personas u objetos?	Director del área de tecnología, administradores y operadores de Base de datos, desarrolladores de aplicaciones.
3. ¿Sobre qué aspectos?	Indicadores (matriz de operacionalización de variables)
4. ¿Quién, quiénes?	Investigadora
5. ¿Cuándo?	Mes de julio 2011 a julio 2012.
6. ¿Dónde?	Empresa VSYSTEMS
7. ¿Cuántas veces?	Dos
8. ¿Qué técnicas de recolección?	Encuesta Entrevista Observación
9. ¿Con qué?	Cuestionario Guía de la Entrevista
10. ¿En qué situación?	Previa cita

Tabla N° 5: Plan de recolección de la Información

Fuente: Investigadora

3.8. Plan de procesamiento de la información

- Revisión crítica de la información recogida; es decir, limpieza de la información defectuosa: contradictoria, incompleta, no pertinente, etc.
- Repetición de la recolección, en ciertos casos individuales, para corregir fallas de contestación.
- Tabulación o cuadros según variables de cada hipótesis.
- Manejo de información (reajuste de cuadros con casillas vacías o con datos tan reducidos cuantitativamente, que no influyen significativamente en los análisis).
- Estudio estadístico de datos para presentación de resultados.
- La presentación de datos puede hacerse siguiendo los siguientes procedimientos:
 - Representación escrita
 - Representación tabular
 - Representación gráfica

CAPITULO IV

4. ANÁLISIS E INTERPRETACIÓN DE RESULTADOS

Para la recolección de la información se aplicó una encuesta de tipo cuestionario cerrado, mediante la formulación de preguntas realizadas a los miembros de la población de manera verbal y registrada por la investigadora.

El objetivo de dicha encuesta fue determinar el conocimiento de metodologías y procesos de migración de datos así como el estado y entorno en el que se encuentra la base de datos.

Como en éste estudio la población es pequeña (N es menor que 100) usamos la totalidad de la descripción de la población.

La población está constituida por el administrador de base de datos de la empresa VSYSTEMS y el personal del área de desarrollo que son usuarios avanzados de base de datos.

4.1. Análisis e interpretación

Pregunta 1:

El proceso de migración que se va a implementar en la empresa es causado por:

OPCIONES	VALOR	PORCENTAJE
Innovación Tecnológica	2	33.33
Ahorro de licencia	4	66.67
El gestor actual no cumple con requerimientos	0	0.00
TOTAL	6	100.00

Tabla N° 6: Tabulación origen del proceso de migración

Fuente: Investigadora

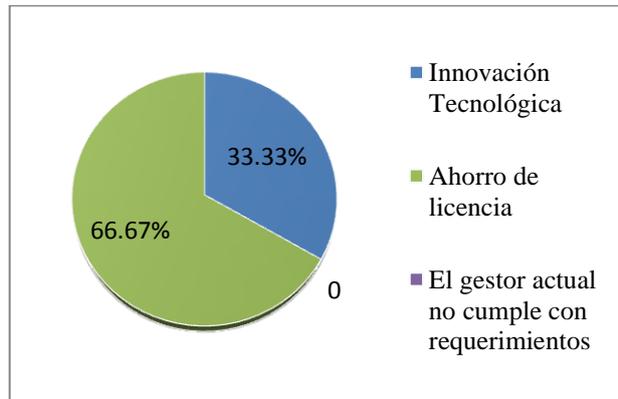


Gráfico N° 8: Tabulación origen del proceso de migración

Fuente: Investigadora.

Interpretación:

De la encuesta realizada se observa que el 66,67% de la población piensa que la principal causa de migración de base de datos es el ahorro de licencia mientras que el 33,33% opina que la migración fue motivada por innovación tecnológica.

De este resultado podemos apreciar que el ahorro de licencias es un factor determinante en la generación de un proceso de migración.

Pregunta 2:

Considera que la estructura de la base de datos de la empresa es:

OPCIONES	VALOR	PORCENTAJE
Muy compleja	2	33.33
Compleja	4	66.67
Simple	0	0.00
TOTAL		100.00

Tabla N° 7: Tabulación complejidad de la base de datos

Fuente: Investigadora

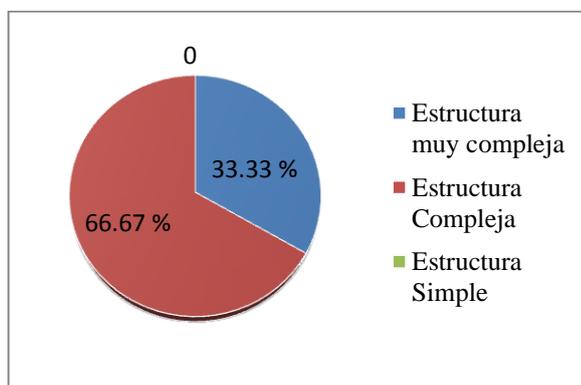


Gráfico N° 9: Tabulación complejidad de la base de datos

Fuente: Investigadora

Interpretación:

De la encuesta realizada se observa que el 67% de la población considera que la base de datos con la que trabaja es de estructura compleja mientras que el 33% opina que la estructura de la base de datos es muy compleja.

De este resultado podemos apreciar que la base de datos que intervendrá en el proceso migratorio de manera generalizada se puede considerar de estructura compleja.

Pregunta 3:

Conoce o disponen de una metodología para aplicar la migración:

OPCIONES	VALOR	PORCENTAJE
SI	2	33.33
NO	4	67.67
TOTAL		100.00

Tabla N° 8: Tabulación de conocimiento de metodologías de migración

Fuente: Investigadora

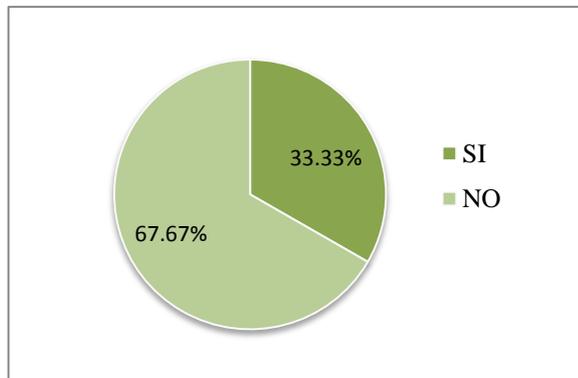


Gráfico N° 10: Tabulación Conocimiento de metodologías de migración

Fuente: Investigadora

Interpretación:

De la encuesta realizada se observa que el 67,67 % de la población no conoce una metodología para migración de datos.

De este resultado podemos concluir que la mayor parte del personal de la empresa VSYSTEMS no conoce o no dispone de una metodología para migración de datos.

Pregunta 4:

El tamaño de la base de datos es:

OPCIONES	VALOR	PORCENTAJE
Grande	0	0.00
Mediano	2	33.33
Pequeño	4	66.67
TOTAL		100.00

Tabla N° 9: Tamaño de la base de datos

Fuente: Investigadora

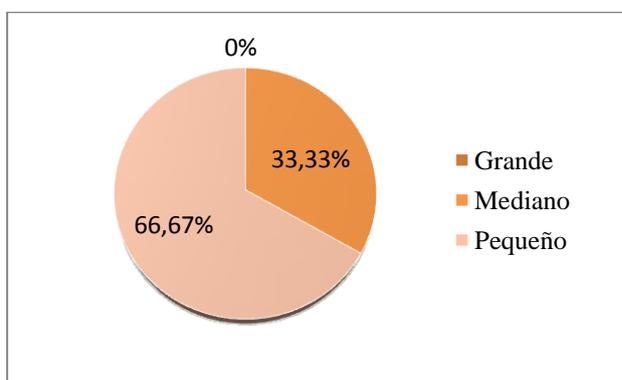


Gráfico N° 11: Tamaño de la base de datos

Fuente: Investigadora

Interpretación:

De la encuesta realizada se observa que el 67% de la población considera que la base de datos con la que trabaja se puede considerar como pequeña (menos de 100 tablas) mientras que el 33% considera que el tamaño de la base de datos es mediano (entre 100 y 999).

De este resultado podemos apreciar que la base de datos que intervendrá en el proceso migratorio es de tamaño pequeña.

Pregunta 5:

Cuál de los siguientes aspectos considera usted que es mayor gravedad al realizar un proceso de migración.

OPCIONES	VALOR	PORCENTAJE
Pérdida de información	4	66.67
Pérdida de integridad	2	33.33
Incompatibilidad de datos	0	0.00
TOTAL		100.00

Tabla N° 10: Tabulación problemas con la información migrada

Fuente: Investigadora

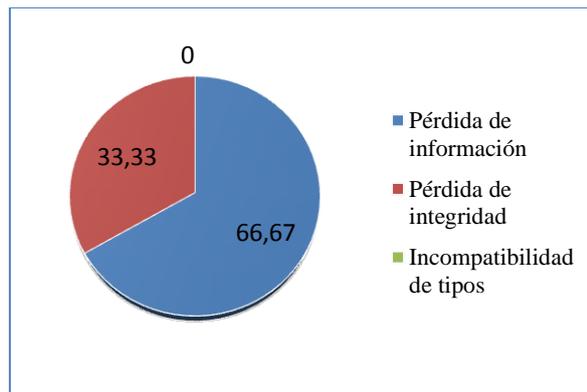


Gráfico N° 12: Problemas con la información

Fuente: Investigadora

Interpretación:

De la encuesta realizada se observa que el 66.67% de la población considera que el problema de mayor gravedad que se podría presentar en un proceso migratorio de datos sería la pérdida de información mientras que el 33.33% considera que el problema más grave que podría sufrir el proceso de migración es la pérdida de integridad.

De este resultado podemos apreciar que la mayor parte de la población encuestada opina que el problema más preocupante sería la pérdida de información en un proceso migratorio.

Pregunta 6:

La base de datos tiene documentación sobre su diseño y estructura

OPCIONES	VALOR	PORCENTAJE
SI	6	100.00
NO	0	100.00
TOTAL		100.00

Tabla N° 11: Tabulación existencia de documentación de base de datos.

Fuente: Investigadora

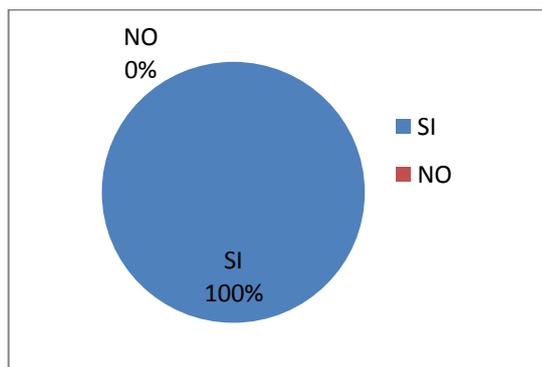


Gráfico N° 13: Existencia de documentación de base de datos

Fuente: Investigadora

Interpretación:

De la encuesta realizada se observa que el 100% de la población dice contar con documentación e información de diseño de la base de datos.

De éste resultado podemos concluir que la base de datos que participará en el proceso migratorio cuenta con documentación necesaria.

Pregunta 7:

Realiza análisis de información y conoce cuál es la información crítica de su base de datos.

OPCIONES	VALOR	PORCENTAJE
SI	0	0.00
NO	6	100.00
TOTAL		100.00

Tabla N° 12: Tabulación respecto a análisis de información realizados.

Fuente: Investigadora

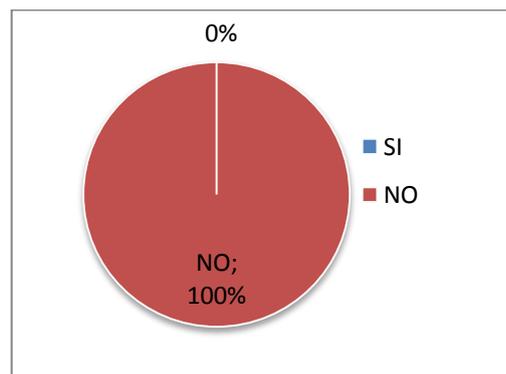


Gráfico N° 14: Tabulación se realiza análisis de información.

Fuente: Investigadora

Interpretación:

De la encuesta realizada se observa que el 100% de la población dice no contar con un análisis de información que permite definir la información crítica.

De éste resultado podemos la base de datos que participará en el proceso migratorio no tiene definido un proceso de análisis ni identificación de información crítica.

Pregunta 8:

Se realiza algún procedimiento periódico para depuración y ejecutar control de calidad sobre los datos

OPCIONES	VALOR	PORCENTAJE
SI	2	33.33
NO	4	66.67
TOTAL		100.00

Tabla N° 13: Tabulación se realiza procedimientos de depuración en la base de datos.

Fuente: Investigadora

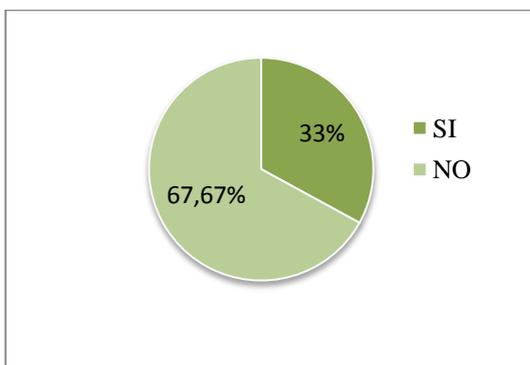


Gráfico N° 15: Se realiza procedimientos de depuración de base de datos.

Fuente: Investigadora

Interpretación:

De la encuesta realizada se observa que el 67.67% de la población dice no realizar control de calidad ni procesos de depuración a la información almacenada en la base de datos.

De este resultado podemos determinar que la información almacenada en la base de datos que participará en el proceso migratorio puede presentar errores pues no se ha realizado una correcta depuración sobre ellos.

Pregunta 9:

Las aplicaciones que interactúan con la base de datos consiste en:

OPCIONES	VALOR	PORCENTAJE
Software desarrollado por la empresa	6	100.00
Software comprado sin personalización	0	0.00
Software comprado con personalización	0	0.00
TOTAL		100.00

Tabla N° 14: Tabulación tipos de aplicaciones que interactúan con la base de datos

Fuente: Investigadora

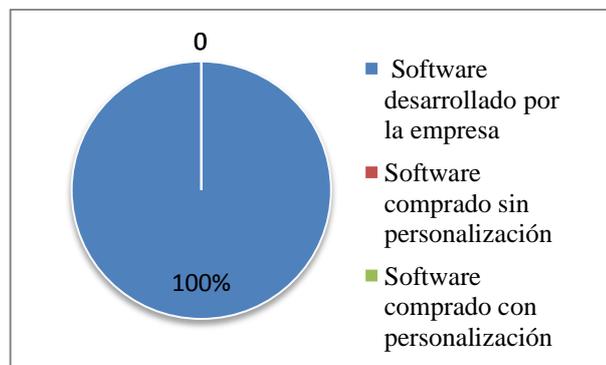


Gráfico N° 16: tipos de aplicaciones que interactúan con la base de datos

Fuente: Investigadora

Interpretación:

De la encuesta realizada se observa que el 100% de la población afirma que la base de datos interactúa con software desarrollado por la empresa VSYSTEMS.

De este resultado podemos observar que la base de datos que participará en el proceso migratorio únicamente se enlaza con aplicaciones desarrolladas por la misma empresa.

Pregunta 10:

El crecimiento de la base de datos en el último año se encuentra dentro de los siguientes porcentajes:

OPCIONES	VALOR	PORCENTAJE
Menos de 25%	4	66.67
Entre 25 y 50%	2	33.33
entre 50 y 75%	0	0.00
Entre 75 y 100%	0	0.00
TOTAL		100.00

Tabla N° 15: Tabulación datos crecimiento de la base de datos

Fuente: Investigadora

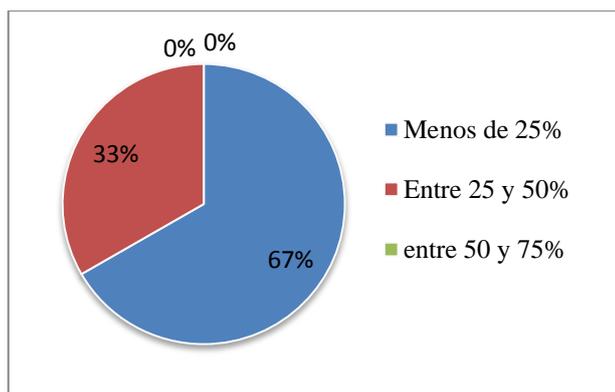


Gráfico N° 17: Crecimiento anual de base de datos.

Fuente: Investigadora

Interpretación:

De la encuesta realizada se observa que el 67.67% de la población afirma que la base de datos ha tenido un crecimiento anual menor a 25%, mientras que el 33.33% de la población afirma que el crecimiento anual se ubica entre el 25% y 50%.

De este resultado podemos observar que la base de datos que participará en el proceso migratorio tiene un crecimiento anual definido mayoritariamente en un porcentaje inferior al 25%.

Pregunta 11:

Luego de haber desarrollado algún proceso migratorio su base de datos presentó inconsistencias graves.

OPCIONES	VALOR	PORCENTAJE
SI	5	33.33
NO	1	66.67
TOTAL		100.00

Tabla N° 16: Tabulación existencia de inconsistencias en procesos migratorios previos.

Fuente: Investigadora

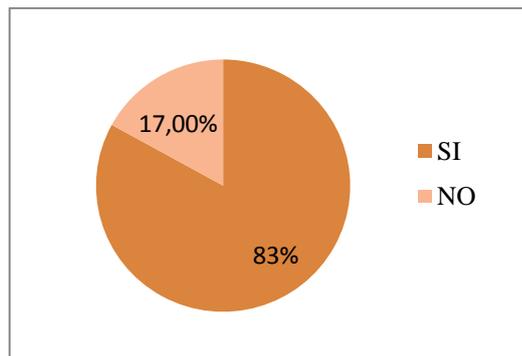


Gráfico N° 18: Existencia de inconsistencias en procesos migratorios previos.

Fuente: Investigadora

Interpretación:

De la encuesta realizada se observa que el 83% de la población afirma que la base de datos ha presentado inconsistencias graves luego de haber pasado por un proceso de migración.

4.2. Verificación de la hipótesis

Al establecer una metodología para aplicar migración de datos se optimizará y asegurará la integridad y consistencia de la información administrada en la empresa VSYSTEMS.

Hipótesis nula (H0):

Al establecer una metodología para aplicar migración de datos NO SE optimizará ni se asegurará la conservación integridad y consistencia de la información administrada por la empresa VSYSTEMS.

Hipótesis alternativa (H1)

Al establecer una metodología para aplicar migración de datos se optimizará y asegurará la integridad y consistencia de la información administrada por la empresa VSYSTEMS.

4.2.1. Modelo Estadístico

Elección de la prueba de hipótesis:

Para verificación de hipótesis se utilizó la prueba de Chi-cuadrado que es una prueba estadística que permite evaluar la hipótesis analizando la independencia de variables.

La fórmula que usa Chi-cuadrado es:

$$x^2 = \frac{\sum(f_o - f_e)}{f_e}$$

Siendo

x^2 chi-cuadrado

f_o frecuencia observada

f_e frecuencia esperada

4.2.2. Definición del nivel de significancia

El nivel de significancia elegido es de 0.05

$$\alpha = 0.05 \%$$

4.2.3. Especificación de la región de aceptación y rechazo

Grados de libertad

$$gl = (\text{renglones} - 1) * (\text{columnas} - 1)$$

El número de renglones de las frecuencias observadas para el presente modelo estadístico es 2 y el número de columnas 2, por lo tanto el grado de libertad es el resultado de las siguientes operaciones:

$$gl = (2 - 1) * (2 - 1)$$

$$gl = (1) * (1)$$

$$gl = 1$$

El valor crítico definido para 1 grado de libertad en la tabla de distribución de chi-cuadrado con un nivel de significancia de 0.05 es: 3. 8414.

4.2.4. Recolección de cálculos y datos estadísticos

La información que servirá para la aplicación de la prueba estadística respecto a las frecuencias observadas se encuentra en la Tabla N° 17.

	SI	NO	TOTAL
Conoce o disponen de una metodología para aplicar la migración	2	4	6
La base de datos ha presentado inconsistencias graves luego de haber pasado por un proceso de migración.	5	1	6
	7	5	12

Tabla N° 17: Frecuencias observadas

Fuente: Investigadora

La frecuencia esperada la podemos definir mediante la siguiente fórmula:

$$f_e = \frac{(total\ marginal\ de\ renglones) * (total\ marginal\ de\ columnas)}{N}$$

La tabla N° 18 contiene los datos usados como frecuencias esperadas.

	SI	NO	TOTAL
Conoce o disponen de una metodología para aplicar la migración	3.5	2.5	3
la base de datos ha presentado inconsistencias graves luego de haber pasado por un proceso de migración.	3.5	2.5	3
	7	5	6

Tabla N° 18 : Frecuencias esperadas

Fuente: Investigadora

4.2.5. Cálculo del Chi cuadrado

Una vez determinada la frecuencia esperada se aplicó la fórmula:

$$x^2 = \frac{\sum(f_o - f_e)}{f_e}$$

La Tabla N° 19 contiene el cálculo de chi-cuadrado del modelo estadístico aplicado en la investigación.

	FO	FE	FO- FE	(FO- FE)2	(FO- FE)2/FE
Conoce o disponen de una metodología para aplicar la migración	2	3.5	-1.5	2.25	0.6429
NO Conoce o disponen de una metodología para aplicar la migración	4	2.5	1.5	2.25	0.9000
la base de datos ha presentado inconsistencias graves luego de haber pasado por un proceso de migración.	5	3.5	1.5	2.25	1.5429
la base de datos NO ha presentado inconsistencias graves luego de haber pasado por un proceso de migración.	1	2.5	-1.5	2.25	2.4429
					5.5286

Tabla N° 19: Cálculo de Chi-cuadrado

Fuente: Investigadora

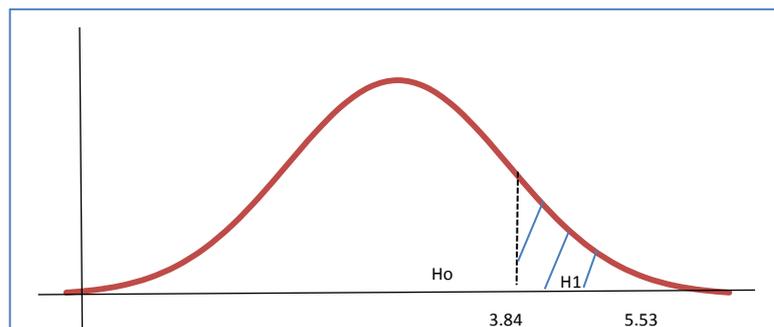


Gráfico N° 19: Verificación de hipótesis

Fuente: Investigadora

Decisión:

El valor $\chi^2 t = 3.84 < \chi^2 c = 5.52$ por lo tanto se rechaza la hipótesis nula y se acepta la hipótesis alternativa.

Análisis:

Al realizar el análisis de la información recolectada en la presente investigación se ha detectado inconformidad en la realización de procesos de migración realizados anteriormente en la empresa VSYSTEMS que ha dificultado el desarrollo de las actividades empresariales y administración de la información en la empresa VSYSTEMS.

La empresa VSYSTEMS no cuenta con una metodología diseñada a medida para realizar el proceso de migración de datos por lo que se necesita desarrollar y aplicar una metodología de migración de base de datos que permita asegurar la consistencia e integridad de los datos.

CAPITULO V

5. CONCLUSIONES Y RECOMENDACIONES

5.1. Conclusiones

- La inexistencia de procedimientos de control de calidad de datos y depuración de información en la empresa VSYSTEMS produce dificultad en los procesos de migración realizados sin metodología específica incrementando los riesgos de inconsistencia.
- La falta de una política de análisis y determinación de la información crítica dentro de la empresa VSYSTEMS, hace necesario que la ejecución de un proceso de migración sea mucho más riguroso y específico para garantizar la preservación de la información de acuerdo a las necesidades de la empresa.
- Entre las principales causas que motivan a la empresa VSYSTEMS a ejecutar un proceso de migración es el ahorro en costos por licenciamiento que causa el empleo de software propietario y la búsqueda de innovación tecnológica.
- El desconocimiento de una metodología por la mayor parte del personal de la empresa VSYSTEMS, produce un retraso en el proyecto global de migración provocando incremento tiempo y por lo tanto de costos económicos.

5.2. Recomendaciones

- Se recomienda que se establezca como política el control de calidad de la información para garantizar la consistencia de la información. Se recomienda la aplicación de una metodología de migración de datos que contenga técnicas que ayuden a garantizar la conservación de la consistencia de información durante un proceso de migración de datos.
- Se recomienda la creación de políticas para llevar un control y administración de la información crítica de la empresa VSYSTEMS.
- El uso de herramientas de software libre para migración ayudaría a la reducción en costos por licenciamiento en la empresa VSYSTEMS, la selección y uso las herramientas de migración debe adaptarse a las necesidades de la empresa.
- Se recomienda aplicar una metodología de migración de datos diseñada a medida, que permita reducir tiempos improductivos durante el proceso de migración.

CAPITULO VI

6. PROPUESTA

6.1. Datos Informativos

La empresa VSYSTEMS de la provincia de Santo Domingo de los Tsáchilas se dedica a desarrollo y venta de software, además de ofrecer consultoría en sistemas informáticos. Esta empresa cuenta con algunas aplicaciones, entre ellas se encuentra el sistema de puntos de venta e inventario que trabaja con una base de datos Oracle10g. Este sistema se encuentra instalado en varias empresas clientes y la base de datos respectiva es administrada por la empresa VSYSTEMS dentro de cada empresa cliente.

El proceso de migración que se va a emprender tiene como origen el gestor de Base de datos Oracle10g y como destino PostgreSQL.

La presente propuesta tiene como beneficiario al personal de la empresa VSYSTEMS, optimizando la ejecución de un proceso importante en la administración de la información de la empresa.

La persona responsable de la ejecución de la propuesta es Susana Caraguay Martínez egresada de la maestría en gestión de base de datos, bajo la tutoría del Ing. Mg. Franklin Mayorga Mayorga en calidad de director de la investigación.

6.2. Antecedentes de la propuesta

Toda la información recopilada en el presente proceso investigativo indica que en la empresa VSYSTEMS existe la necesidad de contar con una metodología específica para realizar la migración de datos.

El diseño y ejecución de una metodología específica para migración de datos permite solucionar algunos de los inconvenientes como:

- Incoherencia en la información que atraviesa un proceso de migración de datos
- Inconsistencia en base de datos
- Reducción de tiempo de depuración y activación de la base de datos

6.3. Justificación

El contar con una metodología para la migración de datos permitirá a la empresa VSYSTEMS:

- Reducir los costos que conlleva un proceso de migración
- Reducir los períodos de inactividad de las aplicaciones que hacen uso de la base de datos, que deben detenerse para realizar la migración.
- Facilidad al realizar el proceso de migración asegurando al mismo tiempo su integridad y consistencia.
- Reutilización de la metodología para implantar nuevos gestores de base de datos, permitiéndoles un crecimiento y actualización tecnológica.

6.4. Objetivos

General

- Diseñar una metodología de migración para la empresa VSYSTEMS que permita la conservación de la integridad y consistencia de la información

Específicos

- Desarrollar una guía técnica para ejecutar el proceso de la migración
- Asegurar la integridad de la información

- Garantizar la consistencia de la información luego de pasar por un proceso de migración.

6.5. Análisis de factibilidad

Factibilidad Operativa

La empresa presta la disposición para permitir el acceso a información necesaria para realizar el presente proyecto.

Se cuenta con la aprobación de la empresa para permitir al personal del área de sistemas participar en el proyecto de migración, así como para hacer uso del equipamiento necesario para realizar la migración.

Se realizará la respectiva capacitación y apoyo a los empleados de la empresa VSYSTEMS durante el proceso de migración de base de datos en un ambiente apropiado.

Factibilidad Técnica

Es técnicamente posible porque se cuenta con la plataforma de hardware que consiste en computadores propios de la empresa así como la interconexión de redes de datos.

Con respecto al software se cuenta con la base de datos origen cuyo gestor es Oracle, la base de datos PostgreSQL que es la base destino y que es una base de datos de software libre con licencia BSD.

Además existe la herramienta de migración de software libre como Ora2Pg que facilita la migración de Oracle a PostgreSQL interviniendo en el proceso de extracción y carga de información.

Factibilidad Económica

Existe un ahorro en cuanto a las herramientas de migración y gestor de base de datos destino, pues son de software libre y licenciamiento GNU y BSD.

Haciendo referencia a la escala de remuneraciones mensuales unificadas de acuerdo a resolución N° MRL-2012-0021 considerando al Administrador de base de datos y al líder de proyecto como servidor público 5 con una remuneración de 1212 dólares el costo hora es de 7.57. Considerando al programador de base de datos como servidor público 2 con una remuneración de 901 dólares, el costo hora del administrador sería de 5.63. La tabla N° 20.

PERSONAL	CANTIDAD	costo/hora	HORAS	TOTAL
Administrador de BD de la empresa	1	7,57	240	1816,8
Programador de base de datos (recurso de la empresa)	2	5,63	240	1351,2
Líder proyecto migración	1	7,57	480	3633,6
TOTAL				6801,6

Tabla N° 20: Gastos empleados en recurso humano.

Fuente: Investigadora.

La tabla N° 21 describe otros gastos adicionales que se emplearon en el presente proyecto:

ITEM	DESCRIPCION	UNIDAD	CANTIDAD	VALOR UNITARIO	VALOR TOTAL
1	Hora de internet	hora	143	0,7	100,1
2	Resma de papel A4	c/u	2	3,9	7,8
3	Útiles y material de oficina	lote	1		24
TOTAL					131,9

Tabla N° 21: Gastos Adicionales

Fuente: Investigadora

6.6. Fundamentación Teórica

6.6.1. Gestor de base de datos Oracle

El Gestor de base de datos es un software que permite crear, mantener y administrar la información de una base de datos.

Oracle es un gestor de base de datos objeto-relacional. En su característica de gestor relacional incorpora estructuras estáticas de datos, conocidas como tablas. En su enfoque orientado a objetos permite trabajar de forma relacional y añadir cualquier mejora orientada a objetos.

El gestor de base de datos Oracle a pesar de poseer características de orientación a objetos, mantiene su núcleo enfocado a lo relacional.

Los elementos que forman parte del servidor Oracle son:

- Instancia de Oracle
- Gestor de almacenamiento

El gráfico N° 20 describe la estructura del servidor de base de datos Oracle.

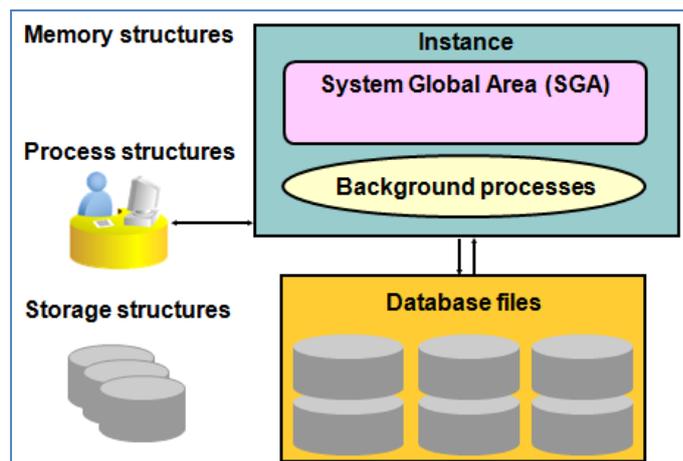


Gráfico N° 20: Estructura de la Base de datos Oracle

Fuente: Azerdark's blog <http://azerdark.wordpress.com/2010/05/12/oracle-database-structure/>

Instancia Oracle

Una instancia de Oracle comprende estructuras de memoria conocidas como SGA (System Global Area) y los procesos background de Oracle. La instancia de base de datos Oracle contiene también los archivos de datos (datafiles), archivos de recuperación de operaciones (redo log files) y los archivos de control (control files). El gráfico N° 21 muestra las estructuras de memoria y procesos para acceder a la base de datos Oracle.

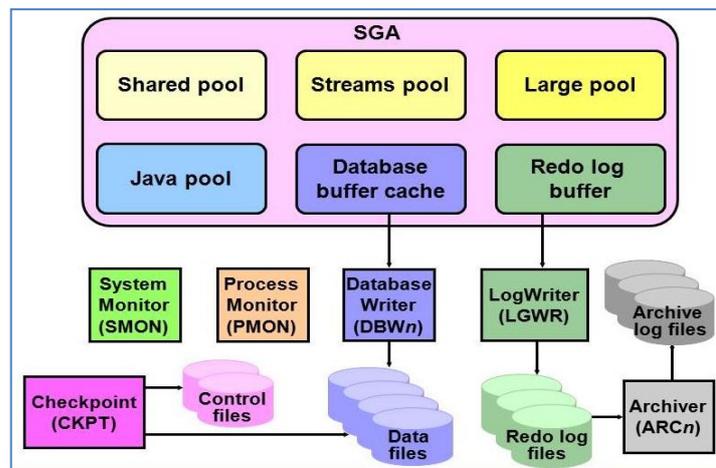


Gráfico N° 21: Instancia Oracle

Fuente: Azerdark's blog <http://azerdark.wordpress.com/2010/05/12/oracle-database-structure/>

Gestor de almacenamiento: El gestor de almacenamiento es el módulo de software que se encarga de interactuar con el gestor de archivos para efectuar operaciones de lectura y escritura sobre el disco del computador. El gestor de almacenamiento de Oracle consta de:

- Estructuras lógicas o internas: Tablespaces (Espacios de tablas), objetos del esquema de usuario.
- Estructuras físicas o externas: Los ficheros de datos almacenados en disco. Los ficheros de datos (asociados a los tablespaces), los ficheros redo log y los ficheros de control.

Las bases de datos relacionales incorporan sólo estructuras estáticas de datos, las llamadas tablas. Esto permite para modelar problemas de datos muy grandes, pero no permite aplicar técnicas de programación orientada a objetos.

Las estructuras internas de Oracle son:

- **Tabla:** Una tabla representa la unidad lógica básica de almacenamiento de información. Se distribuye en un conjunto de filas y columnas. Cada columna tiene un nombre y un tipo de dato asociado a ella.
- **Esquema:** El esquema es una colección de objetos lógicos, utilizados para organizar de manera más comprensible la información, es decir se denomina esquema al conjunto de objetos de una cuenta de usuario.
- **Índices:** Un índice es una estructura de datos cuya función es agilizar la recuperación de los registros, generando búsqueda de una manera más rápida y eficiente.
- **Clúster:** El clúster es un grupo de tablas almacenadas en conjunto físicamente como una sola tabla que comparten una columna en común, esto generalmente se realiza con tablas que son accedidas juntas y frecuentemente, minimizando el número de eventos E/S.
- **Vista:** Una vista es una tabla temporal que surge de la selección de información de una o varias tablas. No pueden ser indexadas ni genera almacenamiento de datos.
- **Secuencias:** Oracle posee un generador de secuencia que se emplea para crear valores enteros secuenciales únicos.
- **Procedimientos y Funciones:** Rutinas escritas en PL/SQL que devuelven un valor escalar o realizan un procedimiento con los datos sin devolución de valor.
- **Triggers:** Los triggers son rutinas que se activan al ejecutar un evento específico sobre una tabla.
- **Sinónimos:** Los sinónimos son objetos del sistema que apuntan a otros objetos. Implementan alias de tablas, vistas, secuencias o unidades de

programas. Por lo general se utilizan para esconder ciertos detalles del objeto que representan al usuario final.

- Segmento de Rollback: Son áreas lógicas de la base de datos que contienen información de las transacciones que se encuentran en curso y que aún no han sido confirmadas o deshechas.
- Espacios de tablas (Tablespaces): Un tablespace es una división lógica de la base de datos. Es un contenedor para tablas, índices y más objetos de base de datos. Cada base de datos tiene al menos uno y un tablespace puede pertenecer sólo a una Base de Datos. Son usados para mantener juntos los datos de usuarios o de aplicaciones para facilitar su mantenimiento o mejorar las prestaciones del sistema.

Dentro de los espacios de tablas (tablespaces) y archivos de datos (datafiles), el espacio utilizado para almacenar datos es controlado por el uso de tres estructuras:

- Segmentos: Un grupo de extensiones utilizados para almacenar un tipo particular de datos. Existen 4 tipos de segmentos: datos, índices, rollback y temporales.
- Extensiones: Es un grupo de bloques de datos. Se establecen en un tamaño fijo y crecen a medida que van almacenando más datos. También se pueden redimensionar para aprovechar mejor el espacio de almacenamiento.
- Bloques: Es la unidad de almacenamiento más pequeña en una base de datos Oracle. Contiene una pequeña porción de información (header) referente al bloque en sí y el resto a los datos que guarda. Generalmente, un bloque de datos ocupará aprox. 2 KB de espacio físico en el disco (asignación típica).

Las estructuras externas de Oracle están contenidas por ficheros que utiliza el servidor de la base de datos. Estos ficheros guardan información tanto de los datos almacenados como de la información necesaria para gestionar la propia base de datos.

Los ficheros correspondientes a la estructura externa de Oracle son:

- Ficheros de BD
- Ficheros redo log
- Ficheros de control
- Ficheros de Traza

Para poder ejecutar SQL sobre la base de datos, hay que conectar con la instancia Oracle de la base de datos, lo cual requiere la comunicación entre un proceso cliente y el servidor. Los componentes utilizados por Oracle para procesar el SQL dependen del código enviado y los eventos que se podrían generar son:

- Las consultas devuelven filas.
- Las instrucciones DML (Lenguaje de Manipulación de Datos) graban cambios.
- La instrucción *commit* asegura el proceso de la transacción.

La conectividad de red con la base de datos Oracle es manejado por un proceso de servidor llamado “Listener” que está configurado para escuchar la conexión en un puerto específico en el servidor de base de datos. Por defecto el puerto de Oracle es 1521.

6.6.2. Gestor de base de datos PostgreSQL

PostgreSQL es un sistema de gestión de bases de datos objeto-relacional (ORDBMS de acuerdo a sus siglas en inglés), con licencia BSD y con su código fuente disponible libremente. Además es un gestor de base de datos multiplataforma.

Originalmente fue desarrollado por un equipo de la Universidad de Berkeley(California) y se denominó POSTGRES, su diseño poseía soporte parcial a la orientación de objetos y conceptos avanzados de base de datos. A partir de

1994 se incluye SQL a Postgres y se libera éste proyecto con muchas mejoras en relación al proyecto original. En 1996 cambia su nombre a PostgreSQL.

Este gestor de base de datos utiliza un modelo cliente/servidor, el programa servidor se llama Postgres y usa multiprocesos para garantizar la estabilidad del sistema.

En un host únicamente se puede ejecutar un único proceso postmaster y se pueden conectar varios clientes vía TCP/IP de manera remota.

Al ser de licencia BSD se permite utilizar, modificar y distribuir a terceros el código fuente y el código binario del programa de software original con o sin modificaciones. Los trabajos derivados pueden optar a licencias de código abierto o comercial. Las distribuciones del software bajo licencia BSD deben incluir copias literales de la licencia, anuncio de copyright y una "negación de responsabilidad".

El gráfico N° 22 describe los principales componentes de PostgreSQL:

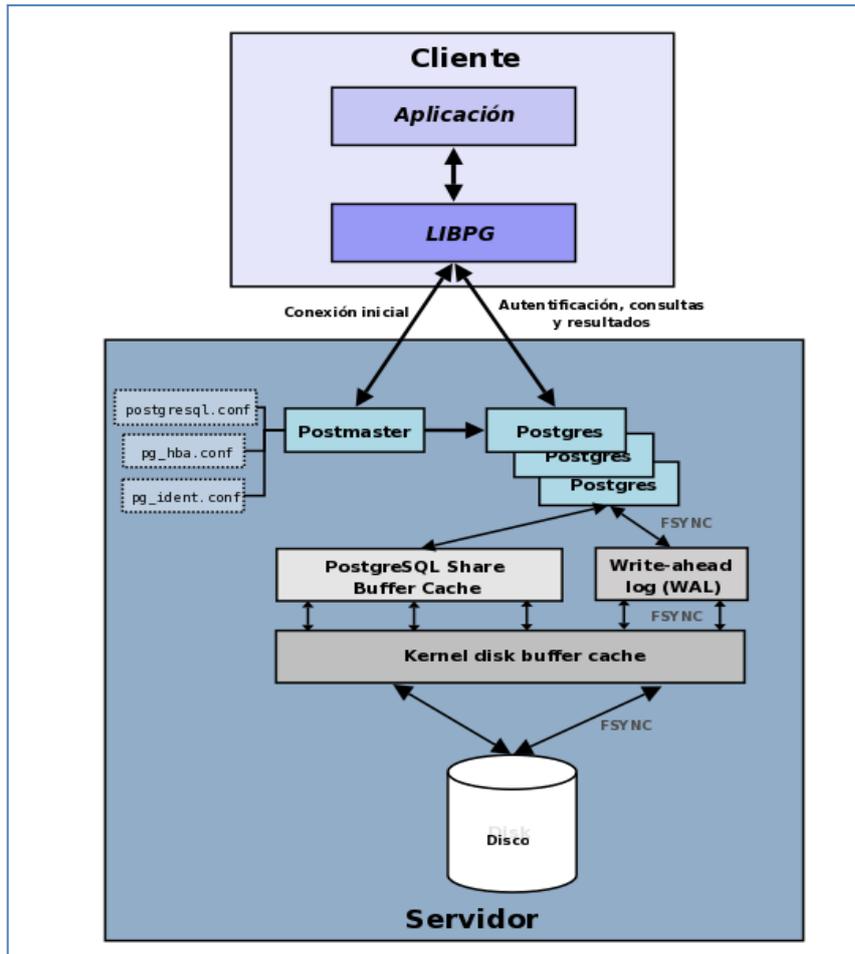


Gráfico N° 22: Principales componentes de postgresQL
Fuente: Portal de la comunidad de usuarios PostgreSQL-es.

Existe un único proceso postmaster y pueden haber varios procesos de servidor postgres. Cada proceso servidor postgres únicamente puede atender a un cliente, es decir pueden existir varios procesos servidor dependiendo del número de clientes que haya. Los componentes internos del servidor de postgresQL son:

- **Aplicación cliente:** Un cliente puede conectarse vía TCP/IP o por sockets, algunos ejemplos de clientes son pgaccess (un cliente que cuenta con interfaz gráfica) y psql (un cliente en modo consola).
- **Demonio postmaster:** Recibe las peticiones de conexión entrante de clientes y es el encargado de crear los procesos hijos para autenticar estas peticiones, gestionar las consultas y mandar los resultados a las aplicaciones clientes.

- Ficheros de configuración: Sirve para especificar parámetros para la base de datos, los tres principales son postgresql.conf, pg_hba.conf y pg_ident.conf.
- PostgreSQL share buffer cache: Memoria compartida usada por PostgreSQL para almacenar datos en caché.
- Write-Ahead Log (WAL): Componente del sistema encargado de asegurar la integridad de los datos (recuperación de tipo REDO).
- Kernel disk buffer cache: Caché de disco del sistema operativo.
- Disco: Disco físico donde se almacenan los datos y toda la información necesaria para que PostgreSQL funcione.

PostgreSQL dispone de versiones para mayoría de los sistemas operativos y cumple totalmente con ACID (del inglés, Atomicity, Consistency, Isolation, Durability) que hacen referencia al procesamiento de transacciones de base de datos que cumplen con las características de atomicidad, consistencia, aislamiento y perdurabilidad. Además tiene soporte para claves foráneas, joins, vistas, disparadores, procedimientos almacenados (en múltiples lenguajes de programación) e integridad transaccional.

El gráfico N° 23 representa un esquema de conexión básico de PostgreSQL.

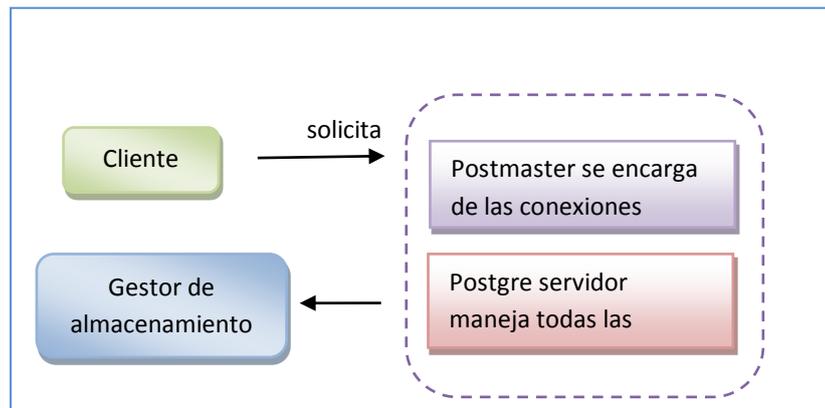


Gráfico N° 23: Esquema de conexión postgresQL

Fuente: Investigadora

PostgreSQL ofrece características tales como:

- Control concurrente multiversión (MVCC) permite el control de concurrencia en su mayor parte sin requerir bloqueos.
- Optimizador de consultas: permite la resolución exitosa de consultas complejas en tiempos adecuados comparables a las bases de datos propietarias.
- Puntos de recuperación en el tiempo (point in time recoveryPITR) que consiste en almacenar y copiar continuamente todas las transacciones de PostgreSQL desde el último backup realizado a nivel de respaldo en disco. Permitiendo la restauración de la base de datos a un determinado punto de tiempo en caso de ser necesario.
- Transacciones anidadas: Mediante uso de savepoints que son marcas internas en el procesamiento de una transacción, se puede deshacer partes específicas de transacción y no toda la transacción por medio del comando ROLLBACK.
- Permite almacenamiento agrupado en tablespaces.
- Posee replicación asíncrona y copias de seguridad en caliente (en línea).

- Es altamente escalable tanto en la cantidad de datos que puede manejar como en el número de usuarios concurrentes que puede atender.

6.6.3. Herramientas de migración

Las herramientas de migración son aplicaciones que facilitan el traslado de estructuras de almacenamiento y datos desde un gestor de base de datos a otro.

Los gestores de base de datos generalmente integran herramientas de migración dentro de sus aplicaciones, pero también existen herramientas de migración desarrolladas por terceros que pueden ser de licenciamiento privativo o de software libre.

Algunas de las herramientas que permiten la migración de base de datos desde Oracle a PostgreSQL son:

EnterpriseDB Migration Toolset

Forma parte de la suite de PostgresPlus Advanced Server, una base de datos desarrollada por la empresa EnterpriseDB que se dedica a ofrecer productos y servicios basados en PostgreSQL.

Postgres Plus Advanced Server cuenta con herramientas propias de migración para transferir datos de Oracle a PostgreSQL que permite migrar todos los datos y la lógica de negocio de la Base de datos Oracle a EnterpriseDB.

Permite realizar migración online en la base de datos destino, realizando creación de objetos inserciones directas sobre la base de datos en el nuevo gestor. También se puede realizar migración offline obteniendo como resultado los esquemas de LDD (Lenguaje de definición de datos) en scripts. El gráfico N° 24 representa los tipos de migración.

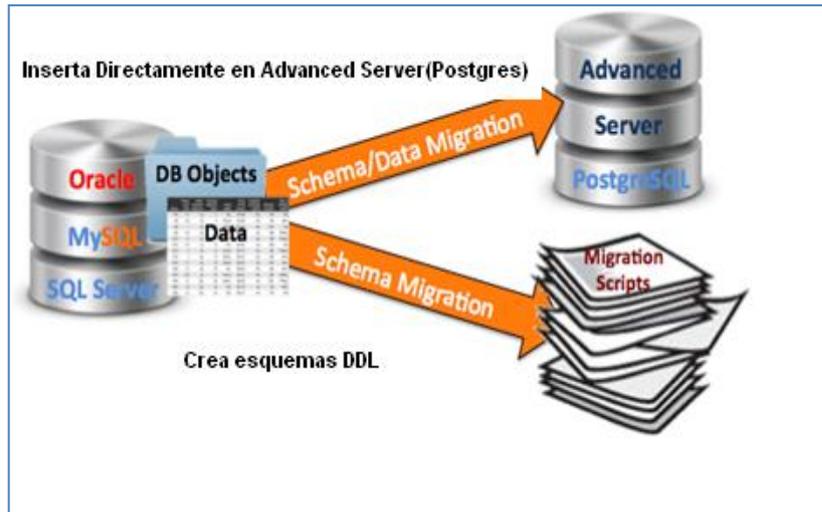


Gráfico N° 24: Tipos de Migración en EnterpriseDB (Online y offline)

Fuente: EnterpriseDB site

SQLWays

Es una herramienta de migración de datos licenciada por InspirerSystems, compatible para migración con algunas base de datos como: DB2, Oracle, MySQL, SQL server.

Ésta herramienta busca automatizar el proceso de migración y facilitar el proceso mediante un entorno de interfaz gráfica y asistentes de migración que permite un uso intuitivo. Permite un filtrado y selección de objetos a ser migrados.

Entre las principales características de ésta herramienta encontramos:

- Migración lógica de negocio del lado del servidor: Convierte procedimientos almacenados, paquetes, funciones y triggers desde Oracle a PostgreSQL.
- Migración de esquemas de base de datos: Conversiones de tablas, índices, vistas, restricciones de integridad y más objetos de base de datos

- Transferencia de datos: Soporta todos los tipos de datos incluidos los de tipo LOB y posee características que le permiten realizar un buen proceso de transferencia de datos.

ORA2pg

Es una herramienta de software libre con licenciamiento GNU, desarrollada como una herramienta multiplataforma para migrar una base de datos Oracle a un esquema de PostgreSQL compatible.

Ora2pg establece una conexión con la base de datos Oracle, extrae la estructura o datos y genera secuencias de comandos SQL en un archivo específico.

Se entorno es ETL, mediante uso de comando de texto y la salida son scripts que contienen sentencias LDD, Los archivos generados pueden ser ejecutados en el gestor de PostgreSQL para la creación de estructuras e inserciones de datos.

Las principales características de Ora2Pg son:

- Exporta el esquema de Oracle a un esquema de PostgreSQL con características compatibles a la versión 7.3 y/o superiores.
- Exporta funciones, desencadenadores, procedimientos, paquetes y cuerpos de paquete predefinidos.
- Exporta datos completos y/o datos filtrados
- Exporta vistas de Oracle y tipos definidos por el usuario
- Proporciona ayuda básica para convertir código PL/SQL a PL/PgSQL (en caso de necesitar depuración manual).
- Exporta privilegios para los usuarios y grupos.

6.6.4. Metodología de migración

Actualmente no existe un estándar específico para procesos de migración, pero existen algunas metodologías de migración que han sido diseñadas a medida para algunas empresas para ejecutar migración de datos.

Algunas de las metodologías de migración existentes son:

- Strategy 7 Corporation: Data Migration and transformation
- Database migration methodology from oracle to sql-server using-ssma de Docserver.
- Migración de datos (Álvaro Alejandro Herrera Jaque).
- The Complete Data Migration Methodology (Joseph R. Hudicka, Dulcian Inc)

6.6.4.1. Metodología de migración Strategy 7 Corporation

Strategy 7 Corporation es una empresa dedicada a ayudar a organizaciones a adaptarse a los cambios tecnológicos. Para cumplir con los objetivos de innovación y cambio tecnológico cuenta con una metodología de migración de datos compuesta por 6 fases.

Las fases que forman parte de ésta metodología son: Análisis, Mapeo, Diseño, Diseño detallado, Construcción, Test & Deploy.

La tabla N° 22 describe las características de la metodología de migración de datos Strategy 7 Corporation.

FASES	ACTIVIDADES
Análisis	Reunir reglas de sistema y negocio Reunir métricas de datos del origen Almacenar información de otros datos(reglas de negocio y origen)
Mapeo	Análisis de mapeo de datos Análisis de integridad Diseño detallado, descripción vías de extracción Determinar medios de transformación Determinar reglas de validación Documentos de requerimientos
Diseño	Estrategia para ejecutar la migración Determinar configuración de software Estrategias para pruebas Estrategias de aseguramiento de calidad Estrategias para certificación
Diseño detallado	Crear proyecto de migración detallado Determinar estándares a usarse Desarrollar especificación de programas Desarrollar planes y benchmark Diseñar funciones, excepciones, manejo, logeos,etc
Construcción	Comunicar estándares de programación Ejecutar motor de migración (herramienta de migración). Extracción, transformación, validación y carga Pruebas de unidad
Test y deploy	Ejecutar plan de test Realizar migraciones simuladas Mejorar reglas de negocio

Tabla N° 22: Características de metodología Strategy 7 Corporation

Fuente: Strategy 7 Corporation

6.6.4.2. Database migration methodology from oracle to sql-server using-ssma

La metodología para migración desde Oracle a SQLServer usando SSMA es una guía para cambio de datos entre gestores Oracle y SQLServer usando herramientas propias de SQLserver y Microsoft. Esta metodología reside en el sitio web DocServer y consta de 4 procesos principales: Estrategia de definición, Análisis, Diseño y Construcción.

La tabla N° 23 muestra las características de la metodología

PROCESO	SUBPROCESO	DESCRIPCION
Estrategia de definición	Inventario	Consolidación de inventario sobre el origen y la base de datos destino, modelos de datos, volúmenes de datos disponible .Uso de Analyzer Migration que es parte de microsoft SSMA toolset.
	Estrategia de mapeo	Definir una estrategia de mapeo, de la fuente al origen. Mediante la herramienta SSMA se puede automatizar el mapeo, pero se debe plantear estrategias para aquellos tipos de datos que no necesitan un manejo distinto
	Selección de herramienta	Propone el uso de SSMA para la migración
	Opciones de puesta en marcha	Opciones para la puesta en marcha en producción, mediante migración incremental o total
Análisis	Análisis de modelos de datos	Analizar la cardinalidad y opcionalidad de los modelos de datos del origen y el destino.
	Análisis de datos	Derivación de los dominios, rangos y valores válidos para los dominios
	Evaluación calidad de datos	Determinación de registros incorrectos en relación a términos de la lógica del negocio
	Extracción, auditoría y criterios de evaluación	Obtener criterios de extracción basados en el análisis de resultados y opciones de implementación. Los criterios de auditoría y validación son basados en los requerimientos para evaluar la correcta y completa migración de datos

Tabla N° 23: Características metodología migración Oracle-sqlServer-SSMA(1)

Fuente: Sitio web Docserver

PROCESO	SUBPROCESO	DESCRIPCION
Diseño	Mapeo	Reglas de mapeo y especificación basada en la estructura de origen y destino
	Diseño de scripts de migración	Secuencia de ejecución de scripts basado en dependencia (técnicas y reglas de negocio)
Construcción	development script	Ejecución de la migración y pruebas de scripts de migración de base de datos, scripts de extracción y scripts de transformación. Auditar y validar la ejecución de scripts

Tabla N° 24: Características metodología migración Oracle-sqlServer-SSMA(2)

Fuente: Sitio web Docserv

6.6.4.3. Metodología de Migración de datos (Álvaro Alejandro Herrera)

Álvaro Herrera en un curso de migración de datos en web denominado Migración de datos, propone una metodología para la migración de datos basada en 9 fases.

La tabla N° 25 presenta las características de la metodología de Herrera.

FASE	ACTIVIDAD
Levantamiento de información	<ol style="list-style-type: none"> 1. Conocer el nuevo modelo de datos. 2. Establecer los conceptos que se deben migrar al nuevo modelo y clasificarlos. 3. Identificar fuentes de los datos. 4. Identificar responsables de los datos. 5. Definición de las tablas críticas a ser migradas. 6. Definición de campos críticos. 7. Identificar ciclos de actualización de la información. 8. Establecer volumen de los datos. 9. Clasificación de datos: grande/mediano/pequeño, o se puede clasificar por rango de datos: 1-30, 31-100, 100-500, etc. 10. Establecer datos históricos que serán traspasados al nuevo sistema: Tratar que se queden en los actuales sistemas. 11. Establecer los flujos de información en el antiguo sistema. 12. Establecer los flujos de información en el nuevo sistema. 13. Establecer la estrategia de migración de datos: 14. Desarrollo de una carga inicial. 15. Desarrollo de una carga incremental. 16. Identificar migraciones de datos que van a ser constantes en el tiempo. 17. Establecer los gap's (brechas) de datos que existirá entre las fuentes de datos v/s el nuevo modelo. Se deberá definir si el gap (brecha) es cubierto por una digitación manual de los datos o un poblamiento automático. 18. Establecer esfuerzo en la corrección de datos antes y después de la limpieza. 19. El entregable de ésta etapa es una planilla con información de objetos 20. Establecer equipos de trabajo de la próxima etapa si es necesario.

Tabla N° 25: Características de metodología de migración Álvaro Herrera (1)

Fuente: <http://www.emagister.com/curso-migracion-datos>

FASE	ACTIVIDAD
Análisis de calidad	<ol style="list-style-type: none"> 1. Establecer una declaración de buenas intenciones con los participantes del proyecto, explicando la metodología a utilizar. 2. Explicar las fuentes de información detectadas en la etapa anterior. 3. Establecer calidad de información en: 4. Integridad 5. Consistencia 6. Nulidad 7. Establecer Herramienta de migración. 8. Establecer forma de migración de cada dato y clasificar en: Manual o Automática. 9. Establecer criterios de calidad de los datos a migrar. 10. Establecer métricas de calidad. 11. Establecer ambientes de desarrollo y pruebas para la construcción de ETL. 12. Coordinar con producción impacto en el desempeño en los actuales productos, al momento de realizar las cargas de datos. 13. Definir y construir herramientas de limpieza automática. 14. Informar calidad de los datos a migrar, completando la siguiente información de la planilla establecida en la etapa de levantamiento de información. 15. Forma de Migración: Manual/Automática 16. Calidad del dato: Buena/Regular/Mala
Transformación / Mapeo	<ol style="list-style-type: none"> 1. Establecer reglas de transformaciones de los datos 2. Establecer secuencia de las migraciones. 3. Definir que cosas se cambiaran en el origen y que cosas son candidatas a transformaciones. 4. Establecer las relaciones entre las fuentes de datos de Origen y de Destino. 5. Establecer esfuerzos de construcción de Extracción de datos, Transformación de datos e Inyección de datos. 6. Establecer si existen agrupaciones de datos desde las fuentes de origen a la de destino o viceversa. 7. Establecer los controles adicionales 8. Definir la administración de errores. 9. Poblar los diccionarios. 10. Tener presente que las reglas pueden sufrir cambios a medida que el proyecto avanza. 11. Diseño del plan de pruebas

Tabla N° 26: Características de metodología de migración Álvaro Herrera (2)

Fuente: <http://www.emagister.com/curso-migracion-datos>

FASE	ACTIVIDAD
Limpeza de datos	<ol style="list-style-type: none"> 1. Establecer criterios de limpieza de los datos. 2. Establecer el esfuerzo real del trabajo de limpieza. 3. Establecer planificación de limpieza de información. 4. Realizar limpieza de datos. 5. Definir herramientas que ayuden a la revisión de la información.
Construcción ETL	<p>Esta etapa, se realiza la construcción de los extractores, transformadores e inyectores de datos.</p> <ol style="list-style-type: none"> 1. Establecer el modelo técnico que se utilizará en los Extractores, Transformadores e Inyectores de datos. 2. Construir herramientas que ayuden a la revisión de información. 3. Construir Extractores de datos. 4. Construir Transformadores de datos. 5. Construir Inyectores de datos al sistema nuevo. 6. Realizar pruebas técnicas unitarias de: Extractores, Transformación, Inyección.
Pruebas	<ol style="list-style-type: none"> 1. Establecer un plan de pruebas para revisar los datos. 2. Por cada regla, se debe definir un escenario. 3. Definir estadísticas de cuadratura. Si existen diferencias establecer claramente los por qué. 4. Establecer criterios de calidad de pruebas. 5. Realizar pruebas funcionales unitarias de: Extractores, Transformación, Inyección. 6. Realizar pruebas de volumen de: Extractores, Transformación, Inyección. 7. Revisar datos cargados. 8. Análisis y clasificación de errores. 9. Revisar los controles adicionales que se deben administrar por restricciones propias del software.
	<ol style="list-style-type: none"> 1. Realizar simulación de cargas reales. 2. Establecer calendario de cargas reales.
Cargas Reales	<ol style="list-style-type: none"> 1. Realizar cargas reales.
Revisión	<ol style="list-style-type: none"> 1. Revisar los datos en el sistema nuevo

Tabla N° 27: Características de metodología de migración Álvaro Herrera (3)

Fuente: <http://www.emagister.com/curso-migracion-datos>

6.6.4.4. The Complete Data Migration Methodology (Joseph R. Hudicka, Dulcian Inc)

Dulcian Inc es una empresa que provee servicios de consultoría, desarrollo de aplicaciones web, datawarehousing y migración de datos específicamente en entorno Oracle.

Joseph Judicka miembro de la empresa Dulcian Inc. propuso una metodología que consta de 7 fases, en base a su experiencia profesional en proyectos de migración.

Las tabla N° 28 muestra las actividades de ésta metodología.

FASES	ACTIVIDADES
Pre estrategia	Se determina el enfoque del proyecto general. Permite determinar el alcance de la migración. Identificar el número de requerimientos de migración de sistemas heredados Conteo de la estructura de datos.
Estrategia	Examinar el plan de migración Evaluar la cantidad de información existente para determinar plazos Elaborar el documento de estrategia de migración, que contiene los objetivos de la migración Calidad de datos de alto nivel
Pre análisis	Se realiza un detalle de los objetivos Se determina responsables de ejecutar las tareas Evaluación de datos de sistemas heredados para determinar si serán trasladados
Análisis	Se obtiene un check list de los datos que serán migrados, esta información tiene tres fuentes principales: Análisis de datos heredados descritos en la fase de pre-análisis Reportes de auditorías. Examinar el uso de los campos Reporte de auditoria de sesiones de usuario
Pre Diseño	Se completa el modelo de datos y se genera el esquema físico de la base de datos
Diseño	Se realiza el mapeo de datos sobre la estructura física creada Generalmente deben participar en esta fase un analista de negocio, un analista de sistemas y un programador/analista que genera las rutinas
Pre-Test / test/ implementación	Permite identificar errores físicos y lógicos Se realiza correcciones de errores físicos Se realiza correcciones de mapeo Ejecutar cargas y evaluar si el número de registros que se carga es el correcto Evaluar si los campos se cargaron correctamente Evaluar si el formato cargado fue el correcto.
Revisión	Todos las modificaciones al modelo de datos, cambios de reglas y scripts modificados son combinados. Se revisa coherencia entre modelo lógico y física.

Tabla N° 28: Características de la metodología de migración de datos de Hudicka

Fuente: Sitio web Dulcian Inc.

6.6.5. Resumen de fases de las metodologías de migración

La tabla N° 29 permite describe la principales fases que contienen las metodologías de migración explicadas en la presente fundamentación teórica.

	STRATEGY 7	DOCSERVER	HERRERA	HUDICKA
FASES	Análisis Mapeo Diseño Diseño detallado Construcción Test & deploy	Estrategia de definición Análisis Diseño Construcción Puesta en marcha Pruebas	Levantamiento de Información Análisis de Calidad Transformación/Mapeo de datos Limpieza de Datos Construcción ETL Pruebas Simulación de Cargas Cargas Reales Revisión de Resultados	Pre Estrategia Estrategia Pre Análisis Análisis Pre-Diseño Diseño Pre-Test/test/ implementación Revisión Mantenimiento

Tabla N° 29: Resumen de fases de metodologías de migración

Fuente: Investigadora

6.6.6. Cuadro comparativo de metodologías de migración

CARACTERISTICA	STRATEGY 7	DOCSEVER	HERRERA	HUDICKA
Análisis proceso migración	X		X	X
Análisis de modelo de datos		X	X	X
Análisis lógica de negocio	X	X		
Estrategia (planificación)	X	X		X
Documentación de requerimientos	X		X	X
Limpieza previa de datos			X	X
Selección de herramienta de migración		X	X	X
Construcción	X	X	X	X
Pruebas	X	X	X	X
Revisión de resultados	X	X	X	X
Análisis de integridad explícito				X
Análisis de consistencia explícito				

Tabla N° 30: Cuadro comparativo de metodologías de migración

Fuente: Investigadora

6.6.7. Análisis de metodologías de migración

Todas las metodologías revisadas incluyen una fase de construcción, pruebas y revisión.

La mayor parte de las metodologías analizadas en la presente fundamentación teórica no consideran una interacción sistémica con el entorno en el que se desenvuelven los datos a ser migrados. Es decir una evaluación del entorno de migración.

No todas las metodologías cuentan con una definición clara y documentada de requerimientos de migración.

Las metodologías de migración no consideran de manera explícita la evaluación de consistencia y la integridad de la información posterior a un proceso de migración.

6.6.8. Metodología de migración de datos propuesta para la empresa

VSYSTEMS

En base a la experiencia adquirida en el presente proyecto se determinó la siguiente metodología de migración de base de datos para la empresa VSYSTEMS.

METODOLOGIA DE MIGRACION DE DATOS PARA LA EMPRESA

VSYSTEMS

INTRODUCCION:

La presente metodología tiene como finalidad proporcionar un instrumento que permita realizar de manera ordenada y sistematizada un proceso de migración de base de datos.

Los objetivos planteados en ésta metodología son:

- Asegurar en cada proceso la conservación de la información.
- Reducir los costos ocasionados por tiempos improductivos de los sistemas que interactúan con las bases de datos que interviene en el proceso de migración.
- Facilitar el traslado de la información entre gestores.
- Estandarizar procesos de migración.

Para cumplir con los objetivos planteados se aplican técnicas en las diferentes fases del proceso para realizar el control de calidad de datos y la integridad de la información. Además se considera los formatos de los resultados entregables al final de los procesos así como fichas para recolección de información en las diferentes actividades de la metodología.

En la presente metodología se ha considerado el uso de partes específicas de ciertos estándares existentes correspondientes a algunas organizaciones internacionales entre ellas tenemos:

- Especificación de *Requisitos* según el estándar de *IEEE 830* (Instituto de Ingenieros Eléctricos y Electrónicos). Ver anexo A1.

NOMENCLATURA

La nomenclatura usada para catalogar a los procesos y actividades está formada por tres partes, la primera letra que identifica si se trata de un proceso o actividad, la segunda parte que hace referencia a la fase y la tercera parte que corresponde al número de secuencia dentro del proceso.

De manera más detallada la nomenclatura empleada es la siguiente:

- Primera letra de la nomenclatura

Para diferenciar los procesos principales de la migración de las actividades, se antepone la letra mayúscula al inicio de las siglas de la siguiente manera:

P identifica que se trata de un proceso

A identifica que se trata de una actividad

- Las letras subsecuentes identifican a la denominación propia del proceso o actividad

En el caso de procesos la segunda letra empleada es M que representa a la palabra migración.

En el caso de actividades la segunda letra corresponde a la identificación del proceso al que corresponde la actividad.

EM Preparación del Entorno de migración

EJ Ejecución de la migración

C Evaluación y control de calidad

- La numeración que se encuentra al final de la nomenclatura corresponde al orden de ejecución de los procesos y actividades

PROCESOS PRINCIPALES:

La metodología consta de tres procesos principales que contienen a su vez actividades y tareas para la migración. Los procesos son:

PM1: Preparación del entorno de migración

AEM1: Planificación de la migración

AEM2: Levantamiento de requerimientos

AEM3: Análisis de procesos de migración

AEM4: Depuración y preparación de la base de datos.

PM2: Ejecución de la migración

AEJ1: Extracción

AEJ2: Definición de estructuras destino

AEJ3: Depuración manual de incompatibilidades

AEJ4: pruebas de carga

AEJ5: Carga de información

PM3: Evaluación y control de calidad

AC1: Análisis de integridad

AC2: Análisis de consistencia

AC3: Evaluación de proceso global de migración

TECNICAS Y PRÁCTICAS

- Especificación de requerimientos basados en Estándar IEEE 830, pero aplicado a gestor de base de datos.
- Pruebas unitarias.
- Pruebas de integridad de datos
- Respaldo de configuraciones
- Catalogación
- Modelado Entidad Relación

El gráfico N° 25 representa la metodología de migración de datos a aplicarse en la empresa VSYSTEMS.

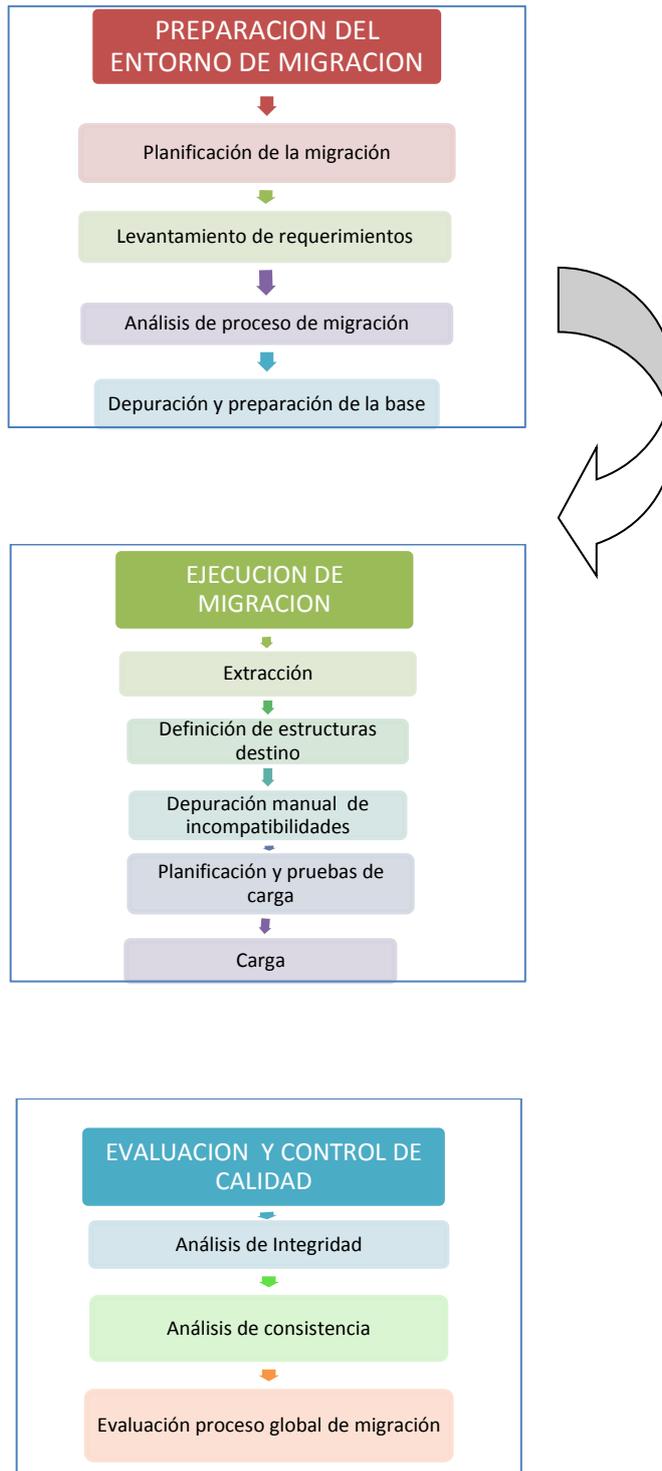


Gráfico N° 25: Metodología de migración

Fuente: Investigadora

PM1: PREPARACION DEL ENTORNO DE MIGRACION

Este proceso se realiza para conocer el estado actual de la base de datos y los sistemas con los que interactúan así como para determinar el aporte de los mismos en el desempeño de las actividades de la empresa.

En ésta fase se incluye algunos aspectos como:

- Conocer la naturaleza de la empresa para determinar lineamientos para la determinación de información crítica
- Determinar plazos y tiempos para la realización de actividades.
- Interactuar con las personas encargadas de la base de datos.
- Identificar roles y funciones del personal encargado de la base de datos.
- Diseñar procesos adecuados para tratar la información
- Preparar la base de datos para la migración.

Las técnicas empleadas en este proceso son:

- Catalogación
- Observación
- Modelado Entidad/Relación

Recurso Humano:

- Personal del área de sistemas de la empresa VSYSTEMS
- Investigadora
- Gerente empresa VSYSTEMS

El gráfico N° 26 demuestra la secuencia de actividades en el proceso de preparación del entorno de migración.

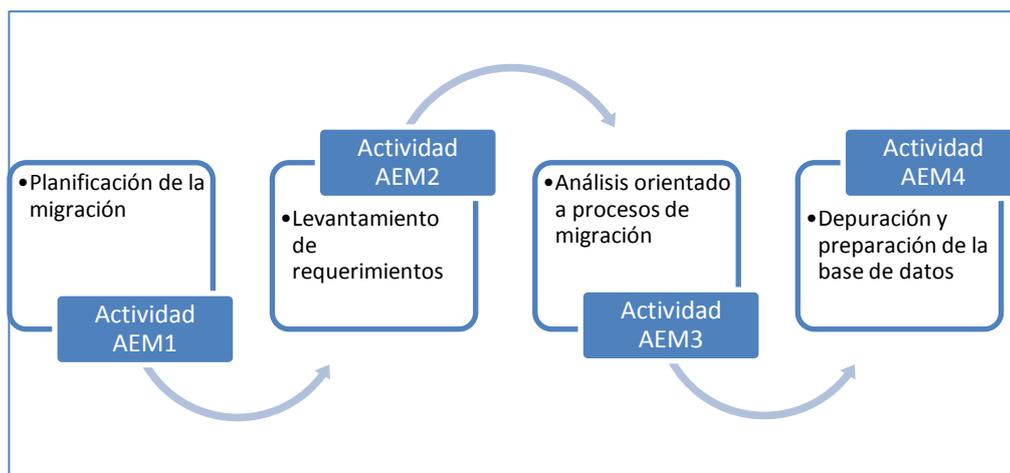


Gráfico N° 26: Proceso de preparación de entorno de migración.

Fuente: Investigadora

AEM1: Planificación de la migración

AEM2: Levantamiento de requerimientos

AEM3: Análisis de procesos de migración

AEM4: Depuración y preparación de la base de datos.

AEM1: Planificación de la migración

La planificación preliminar es indispensable para realizar el proceso de migración ya que permite determinar los tiempos y tareas a ejecutar basándonos en las condiciones generales existentes y características propias de la empresa para realizar el proceso de migración.

Este subproceso está relacionado con el establecimiento de tiempos y metas para el proceso de migración. Se procede a elaborar un plan de migración que contendrá tareas, fechas, responsables e indicadores.

La tabla N° 31 contiene el formato que se usó en la presente metodología de migración para la actividad de planificación preliminar.

OBJETIVOS	META	INDICADOR	ACTIVIDADES	RESULTADOS	PLAZOS

Tabla N° 31: Cuadro de planificación.

Fuente: Investigadora

En la primera columna del formato denominada objetivos, se describen los objetivos que se intenta alcanzar en cada etapa de la migración de datos.

La columna denominada metas corresponde a los objetivos a corto plazo planteados. La columna de indicadores describe la manera como se medirá los logros de cada meta.

En actividades se detalla las tareas a realizar para alcanzar cada meta. En la columna resultados se señala los resultados obtenidos al final de cada actividad.

La columna de plazos corresponde a los tiempos que se emplea en realizar cada actividad.

AEM2: Levantamiento de requerimientos

En ésta actividad se establece los requerimientos generales y específicos que deben considerarse al realizar la migración.

El levantamiento de información consiste en tener un acercamiento con los usuarios del sistema y el administrador de base de datos, se debe recolectar información para su posterior análisis y determinar las características de la migración a realizarse.

Dentro del levantamiento de requerimientos se debe considerar los siguientes aspectos:

Conocimiento del entorno actual.

- Determinar la misión general del sistema de información, es decir la lógica del negocio de la empresa y base de datos que interviene en el proceso de migración.
- Conocer el entorno inicial de la base de datos que se va a migrar, es decir el número de usuarios que hacen uso de la base de datos, expectativa de crecimiento tanto de usuarios como de información
- Determinar la respectiva versión de gestor de base de datos que se está utilizando.
- Conocer la distribución de almacenamiento físico en donde debe almacenarse la información
- Determinar sobre qué equipo funciona el actual gestor de base de datos y si la estructura interna es cliente/servidor o host a host, si son procesos compartidos o sesiones individuales.
- Solicitar manuales de seguridad y procesos de base de datos para determinar si usa alguna estandarización específica en la definición de tablas
- Solicitar logs para determinar que procesos de la base de datos actual presentan mayor inconveniente para su depuración previa a la migración
- Solicitar y analizar el esquema conceptual de la base de datos.
- Analizar el modelo Entidad/relación.

Realizar un análisis de la funcionalidad del sistema (front-end) con el gestor de base de datos a ser migrado.

En ésta tarea se debe solicitar información (diagramas de procesos, casos de uso) en caso de ser necesario para aclarar la interacción de los datos con la aplicación. Si no existiera se debe realizar un análisis identificando cada proceso de interacción con la aplicación

.Los procedimientos a realizar para recolectar la información necesaria son:

- Identificación de tablas que intervienen en el proceso de migración.

Además se debe identificar las tablas que intervienen para determinar las características de la transacción (lectura y/o escritura) y la frecuencia de acceso de dicha transacción. Esto ayudará posteriormente determinar la criticidad de la información. Los datos que se puede recolectar sobre las tablas de la base de datos están descritos en la tabla N° 32.

Nombre
Tipo de transacción (lectura/escritura)
Frecuencia de acceso
Grado de la tabla:
Cardinalidad:
Restricciones de integridad:
Tablas relacionadas:
Procesos de aplicación relacionados:
Observaciones:

Tabla N° 32: Ficha recolección información de tablas

Fuente: Investigadora

- Identificar funciones y procedimientos que intervienen.

La tabla N° 33 contiene los datos que se deben recolectar de las funciones, mientras que la tabla N° 34 señala la información a recolectar de los procedimientos de la base de datos ha ser migrada.

Nombre de la función:
Frecuencia de acceso:
Tipo de dato retornado:
Procesos de aplicación relacionados:

Tabla N° 33: Ficha recolección de información de funciones

Fuente: Investigadora

Nombre del procedimiento: Tipo de acceso: Frecuencia de acceso: Procesos de aplic. relacionados:

Tabla N° 34: Ficha recolección de información de procedimientos

Fuente: Investigadora

Interactuar en el entorno del usuario

En ésta tarea se emplea la técnica de la observación para encontrar procesos que se realicen y no estén documentados, o hayan sido omitidos en los manuales operativos.

Resultado entregable documento de especificación de requerimientos (DER).

Al culminar la actividad de levantamiento de requerimientos se debe generar un documento de especificación de requerimientos que indique la viabilidad de los requerimientos respecto a plazos y recursos, así como el alcance de los requerimientos diferenciando requisitos de rendimiento, de implementación y de funcionalidad.

Además el documento de especificación de requerimientos contiene información de la problemática a solucionar, descripción general de los requerimientos y requerimientos específicos a ser cumplidos.

Las fuentes de información empleadas para levantamiento de requerimientos son:

- Documentación técnica existente en la empresa.
- Entrevistas
- Observación

El personal involucrado en el levantamiento de requerimientos es:

- Administrador de base de datos de la empresa VSYSTEMS

- Usuarios avanzados de base de datos
- Usuarios finales de base de datos
- Investigadora

AEM3: Análisis de procesos de migración de datos

Una vez reunida la información necesaria sobre el proceso de migración de datos a realizarse se analiza la información para definir las acciones a tomar en la siguiente fase. Las actividades a seguir para realizar el análisis de los procesos de migración de datos son:

- Analizar las características entre gestores para predecir posibles puntos de coincidencia o fallos.

Para esto es recomendable instalar de manera autónoma el gestor origen y también el gestor destino para familiarizarse e identificar diferencias y similitudes entre ellos.

- Determinar la herramienta de migración a usar

Evaluar las herramientas disponibles para migración y seleccionar la que más se adapte al criterio y necesidades de la empresa.

Los criterios para realizar la evaluación de herramientas pueden ser:

- Facilidad de uso
- Disponibilidad de soporte
- Costo de licenciamiento.

En ésta actividad es preferible realizar una instalación y pruebas de migración básicas con las herramientas para determinar ventajas y desventajas de cada una de ellas.

- Identificación de información crítica

Se debe analizar la información recogida en la fase de levantamiento de requerimientos, para ello primero se debe identificar la naturaleza del sistema que hace uso de la base de datos.

Luego se identifica las tablas que contienen la principal información sobre la naturaleza de la empresa. Además catalogaremos la importancia de cada tabla comparando contra el número de accesos.

Técnica empleada

La técnica que se emplea en éste proceso es la catalogación. Que consiste en la clasificación de información respecto a los objetos de base de datos.

AEM4: Depuración y preparación de la base de datos.

Durante ésta actividad se realiza la depuración de la información para prevenir errores durante la extracción y carga. Además permite que la migración sea únicamente de información válida y correctamente relacionada.

Para realizar la depuración y preparación de la base de datos a ser migrada es necesaria la ejecución de las siguientes tareas:

- Eliminación o corrección de registros que no son válidos
- Verificar la información para asegurar que la información se encuentre correctamente relacionada y definida en el tipo de dato correspondiente antes de iniciar con la ejecución de la migración.
- Identificar triggers, funciones, procedimientos y más objetos de base de datos que se encuentren deshabilitados para que no formen parte del proceso de migración.
- En una base de datos que posea la auditoría habilitada, se debe establecer el rango de fecha a partir de la cual será migrada la información de auditoría con el fin de no cargar con información innecesaria.

Para facilitar éste proceso se puede usar un registro con la información de objetos e información que no debe migrarse. Es importante documentar éste paso como constancia y autorización de no migración de información innecesaria. La Tabla N° 35 contiene la ficha para registro de las depuraciones realizadas sobre objetos de base de datos.

FECHA	TIPO DE OBJETO	NOMBRE	DESCRIPCION	RESPONSABLE	OBSERVACIONES

Tabla N° 35: Ficha depuración de objetos de base de datos.

Fuente: Investigadora

PM2: EJECUCION DEL PROCESO DE MIGRACION

Consiste en extraer la información del gestor origen, se transforma a lenguaje LDD (Lenguaje de definición de datos) de acuerdo al gestor destino y se realiza la creación de estructuras en el gestor destino. Además se realiza la carga de la información luego de la ejecución de pruebas que garanticen un correcto enlace de datos. En ésta actividad se emplea una herramienta de migración que facilite el proceso de migración.

Las técnicas empleadas en éste proceso son:

- Respaldo de configuraciones
- Modelado Entidad/Relación
- Pruebas unitarias de carga
- Pruebas de integridad de datos

Recurso Humano:

- Personal del área de sistemas de la empresa VSYSTEMS
- Investigadora

La secuencia de actividades en éste proceso es:

AEJ1: Extracción

AEJ2: Definición de estructuras destino

AEJ3: Depuración manual de incompatibilidades

AEJ4: Pruebas de carga

AEJ5: Carga de información

El gráfico N° 27 describe la secuencia de actividades a realizarse en la ejecución del proceso de migración.



Gráfico N° 27: Proceso de ejecución de la migración

Fuente: Investigadora.

AEJ1: Extracción de los datos:

La extracción de datos se puede realizar usando una herramienta de extracción o mediante lenguaje de consulta, que permite extraer la estructura y los datos de la base de datos origen.

Cuando se usa una herramienta de migración, ésta se conecta como cliente a la base de datos para realizar la extracción de las estructuras y/o datos. Con la información extraída internamente la herramienta de migración realiza una conversión de los tipos de datos, de manera que sean compatibles a los tipos de datos del gestor de base de datos destino.

Los objetos de base de datos incluyen: Tablas, vistas, procedimientos, funciones, triggers, secuencias, paquetes y restricciones de integridad.

Se debe asegurar que la extracción de los datos no cause bloqueos de tablas o reducción de rendimiento en el gestor origen.

AEJ2: Definición de estructura destino

Se realiza la creación de la estructura de la base de datos mediante lenguaje DDL en el gestor de destino. Esta creación de estructuras puede ser realizada directamente sobre el gestor destino (migración online) mediante herramientas de migración o puede ser una migración offline en la que se debe ejecutar los Scripts o archivos de ejecución por lotes creados por las herramientas de migración

AEJ3: Depuración manual de incompatibilidades.

En caso que exista alguna incompatibilidad de los tipos de datos entre los dos gestores que la herramienta de migración no pudo solucionar se procederá a realizar el mapeo de tipos manualmente mediante lenguaje LDD.

Para ésto se revisa las estadísticas de la migración, mediante un conteo de objetos válidos e inválidos luego de ejecutar la creación de objetos en el gestor destino y se realiza corrección semántica en la definición de objetos de base de datos.

Técnicas:

- Revisión de estadísticas de migración
- Revisión de fichas de información de tablas, funciones y procedimientos.
- Comparar la semántica y diferencias entre los tipos de datos de las tablas.

AEJ4: Planificación y pruebas de carga

Se establecerá una planificación para ejecutar las respectivas pruebas de carga con el fin de identificar y predecir posibles inconvenientes que podrían suscitarse en la carga definitiva.

El formato señalado en la tabla N° 27 fue usado para llevar un registro de ejecución de pruebas de carga de datos migrados.

FECHA	TIPO DE PRUEBA	ALCANCE	CARDINALIDAD ORIGEN	CARDINALIDAD DESTINO	% EXITO	OBSE RV.

Tabla N° 36: Cuadro planificación de pruebas de carga

Fuente: Investigadora

Los tipos de carga que se puede ejecutar para pruebas de migración son: cargas parciales (pruebas unitarias) y cargas totales.

Pruebas unitarias:

Las pruebas unitarias nos permiten evaluar el funcionamiento de una parte específica de un proceso de migración.

Objetivo: Evaluar la carga de información de tablas que intervienen en ciertos procesos de la actividad de la empresa.

Descripción: Busca determinar la completitud de la información migrada. Es decir una total exactitud entre las filas de datos migradas y la cantidad de registros existentes en la base de datos destino, analizando objetos específicos de la base de datos.

Técnica: Comparar resultado esperado con el resultado obtenido.

Reportar errores.

Pruebas de integridad de datos

Las pruebas de integridad permiten evaluar que la integridad sea conservada durante un proceso de migración de datos.

Objetivo: Evaluar el cumplimiento de integridad

Descripción: Se debe realizar una verificación de que las reglas de integridad fueron correctamente migradas.

Técnica:

Verificación de existencia de la regla de integridad

Verificación semántica de la regla de integridad.

Realizar inserciones, eliminación y actualizaciones manuales sobre la base de datos destino con datos válidos y datos inválidos para verificar cumplimiento de integridad.

Una vez realizadas las pruebas se analiza los resultados y se clasifica los tipos de errores presentados para su posterior depuración. Se realizará las modificaciones a nivel de scripts para su ejecución y corrección.

AEJ5: Carga de la información

Una vez determinada la estructura se procede a cargar la información en las tablas de la base de datos destino.

Se debe considerar los límites en los que se encuentran las secuencias para que los nuevos datos sean ingresados con las secuencias sincronizadas.

PM3: EVALUACION Y CONTROL DE CALIDAD

Una vez realizada la carga de información se revisa los objetos e información de la base de datos destino de la migración, así se determina si se están cumpliendo las reglas de integridad, la consistencia y validez de la información. Además permite evaluar el cumplimiento de requerimientos planteados por la empresa.

Las técnicas empleadas son:

- Respaldo de configuraciones
- Modelado Entidad/Relación
- Pruebas de integridad de datos

Recurso Humano:

- Personal del área de sistemas de la empresa VSYSTEMS
- Investigadora

La secuencia de actividades en éste proceso es:

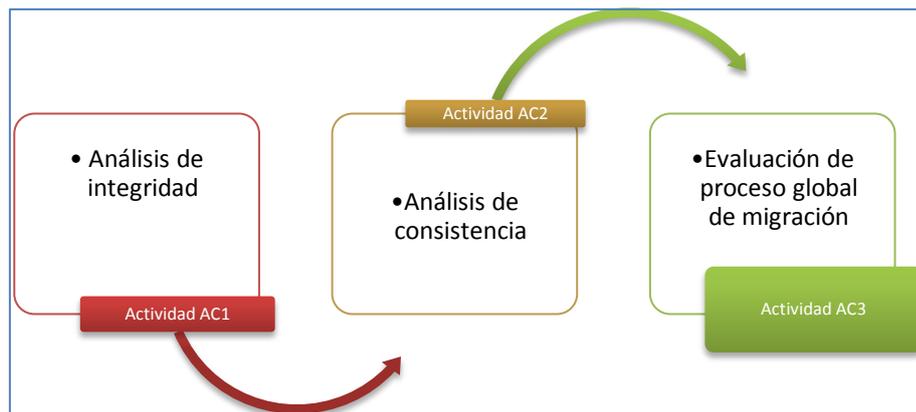


Gráfico N° 28: Proceso de evaluación y control de calidad

Fuente: Investigadora

AC1: Análisis de integridad

AC2: Análisis de consistencia

AC3: Evaluación de proceso global de migración

AC1: Análisis de integridad

Además de la evaluación de las pruebas de integridad realizada en la fase de pruebas de carga se ejecuta el análisis de integridad mediante pruebas de integridad total sobre la base de datos migrada usando como técnica inserciones, eliminaciones y actualizaciones manuales sobre las tablas de la base de datos destino con dato válidos y datos inválidos para verificar cumplimiento de integridad, asignando cuantificación al cumplimiento o no cumplimiento de las mismas.

Usualmente se puede generar scripts con las transacciones que se ejecutarán y serán evaluadas posteriormente. Al final obtendremos un valor % sobre el cumplimiento de la integridad.

AC2: Análisis de consistencia

Para la comprobación de consistencia se ejecuta una revisión de datos de las tablas críticas para determinar el nivel de exactitud e idoneidad de los datos.

Para evaluación de funciones y procedimientos efectuaremos pruebas de caja negra sobre ellas.

Pruebas de caja negra

Objetivo: se evalúa a nivel de entradas y salidas que produce sin considerar su estructura interna

Descripción: Evaluaremos las funciones migradas para determinar si sus salidas son correctas.

Técnica: Comparar el resultado esperado con el resultado obtenido.

Si existen errores, reportarlos.

AC3: Evaluación de proceso global de migración

Una vez realizado el proceso de migración se evalúa el nivel de efectividad del proceso y de ser necesario se realiza afinamiento de la base de datos y depuración de información.

TECNICAS Y PRÁCTICAS

Una técnica es un conjunto de saberes prácticos o procedimientos que requieren destreza manual y/o intelectual para lograr un resultado específico.

Desde el enfoque de ingeniería en sistemas una técnica son procedimientos que permiten la administración de información mediante el uso de computadores, para lograr un resultado determinado en un sistema de información.

Entre las técnicas que se emplearon en la presente metodología de migración son:

- **Especificación de requerimientos basados en Estándar IEEE 830:**(Anexo A1): Propuesto por el Instituto de Ingenieros Eléctricos y Electrónicos en el año 1998 para estandarizar procesos de definición de requerimientos de software. En el presente proyecto se adaptó ciertos aspectos del estándar adaptándolo a las necesidades de un proceso de migración.
- **Pruebas unitarias:** Son formas de probar el correcto funcionamiento de una parte determinada de un sistema, en el caso de éste proyecto probar ciertas funcionalidades, tablas y más objetos de base de datos asociados a un módulo específico del sistema que interacciona con la base de datos migrada.

Objetivo: Evaluar la carga de información de tablas que intervienen en ciertos procesos de la actividad de la empresa.

Descripción: Busca determinar la completitud de la información migrada. Es decir una total exactitud entre las filas de datos migradas y la cantidad de registros existentes en la base de datos destino, analizando objetos específicos de la base de datos.

Técnica: Comparar resultado esperado con el resultado obtenido.

Reportar errores.

- **Pruebas de integridad de datos:** En el presente proyecto son pruebas que permiten verificar la integridad de las bases de datos, mediante inserciones, actualización y verificación de cumplimiento de reglas

Objetivo: Evaluar el cumplimiento de integridad

Descripción: Se debe realizar una verificación de que las reglas de integridad fueron correctamente migradas.

Técnica:

Verificación de existencia de la regla de integridad

Verificación semántica de la regla de integridad.

- **Respaldo de configuraciones**

Un respaldo de configuraciones consiste en sacar una copia de seguridad de los archivos, scripts o cualquier objeto que contenga información de la configuración de una herramienta determinada. En la presente metodología sería específicamente de los respaldos de la configuración de la Base de datos PostgreSQL que es la base destino de la migración, incluido parámetros propios del sistema operativo como son las variables de entorno.

Además de la información que se encuentra contenida en la misma.

Los respaldos deben ser confiables y minimizar las posibilidades de error. El soporte de almacenamiento del respaldo debe ser almacenado en un lugar seguro y diferente al host donde se encuentra la base de datos y debe permitir una recuperación rápida y eficiente.

El detalle del respaldo de la configuración se encuentra en el Anexo A3.

- **Catalogación:** Técnica que consiste en clasificar la información luego de realizar un análisis. En la presente metodología la catalogación se refiere a clasificar en base a criterios los diferentes objetos de base de datos, los criterios pueden ser criticidad de la información, numero de accesos o procesos de aplicación en los que intervienen.

- **Modelado Entidad Relación:** La diagramación de modelos entidad/relación permite tener una visión de las relaciones entre los datos. La revisión de los modelos permite ir controlando la conservación de dichas relaciones y determinando pautas para el proceso de migración. Anexo A6. Modelado de datos.

6.7. Metodología. Modelo Operativo

6.7.1. Modelo Operativo

El gráfico N° 29 demuestra el modelo operativo de la ejecución de la metodología de migración de base de datos en la empresa VSYSTEMS.

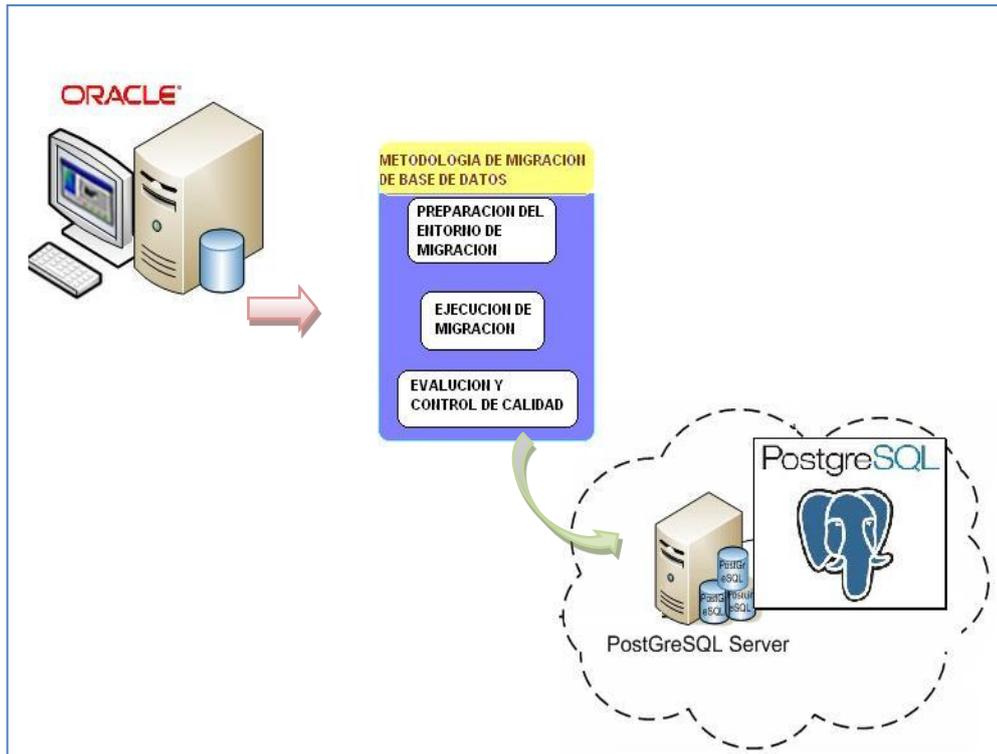


Gráfico N° 29: Modelo Operativo

Fuente: Investigadora

6.7.2. APLICACION DE LA METODOLOGIA DE MIGRACION DE BASE DE DATOS EN LA EMPRESA VSYSTEMS

La aplicación de la metodología de migración de datos en la empresa VSYSTEMS permitió la obtención de los siguientes resultados en cada etapa:

6.7.2.1. PM1: Preparación del Entorno de migración.

AEM1: Planificación

La Tabla N° 37 contiene la planificación realizada para ejecutar la metodología de migración de datos en la empresa VSYSTEMS.

OBJETIVO	META	INDICADOR	ACTIVIDADES	RESULTADOS	PLAZOS
Preparar el entorno de migración	Establecer requerimientos y características de la migración	Número de requerimientos	Realizar entrevistas y encuestas.	Documento de requerimientos	5 días
	Analizar similitudes y diferencias entre gestores de base de datos		Instalar gestores para su comparación. Revisar documentación bibliográfica y artículos técnicos	Cuadro comparativo de características	10 días
	Determinar herramientas de migración a usar.		Revisión de artículos técnicos. Instalación de herramienta de migración.	Herramientas de migración y manipulación instaladas	10 días
	Identificar información crítica	Número de tablas críticas	Realizar análisis de información y base de datos	Catalogación de tablas por criticidad	15 días
	Preparar la información a ser migrada	Número de registros depurados	Analizar información y realizar documentación. Seleccionar herramienta de migración	Base de datos depurada	15 días

Tabla N° 37: Planificación de la migración (1)

Fuente: Investigadora

OBJETIVO	META	INDICADOR	ACTIVIDADES	RESULTADOS	PLAZOS
Ejecutar el proceso de migración	Extracción de los datos	número de objetos de base de datos a ser migrados	Instalar y usar herramienta de migración para extraer estructura y los datos	Script u objeto contenedor de información	4 días
	Definición de estructuras destino	número de objetos de base de datos creados	Ejecutar sentencias DML para el nuevo gestor / hacer uso de herramienta de migración creando directamente sobre el nuevo gestor	Estructura de base de datos creada en el nuevo gestor	4 días
	Depuración manual de incompatibilidades	número de ocurrencias de incompatibilidad	Analizar la nueva base de datos y establecer diferencias lógicas con la base de datos origen	Estructura de base depurada	15 días
	Realizar pruebas de migración	% de efectividad de pruebas	Realizar pruebas de mapeo e integridad	Documento de registro de pruebas	15 días
	carga de información	% de efectividad de carga	Mapear información a la nueva base de datos	Base de datos migrada	2 días
Evaluar y realizar control de calidad	Determinar eficacia de migración y asegurar calidad de información	% efectividad de proceso de migración	Evaluar estructuras e información para asegurar que sea correspondiente a la información de origen	Base de datos consistente	15 días

Tabla N° 38: Planificación de la migración (2)

Fuente: Investigadora

AEM2: Levantamiento de requerimientos

Se realizó el levantamiento de información en la empresa VSYSTEMS, se obtuvo como resultado el documento de requerimientos de migración para la empresa VSYSTEMS. (Ver Anexo 2).

AEM3: Análisis de procesos de migración

Características de los gestores que intervienen en la migración

Se realizó la instalación de los dos gestores para analizar sus características básicas, además de buscar información en medios bibliográficos y artículos técnicos.

Las características de los gestores se detallan en las tablas N°39 a la N°43.

CARACTERISTICA	DESCRIPCION	POSTGRESQL	ORACLE
Gestor de Base de datos	El Sistema de Gestión de Bases de Datos (SGBD) consiste en un conjunto de programas, procedimientos y lenguajes que nos proporcionan las herramientas necesarias para trabajar con una base de datos.	Es un sistema objeto-relacional, incluye características de la orientación a objetos, como: herencia, tipos de datos, funciones, restricciones, disparadores, reglas e integridad transaccional. A pesar de esto, no es un sistema de gestión de bases de datos puramente orientado a objetos.	Es un sistema de gestión de base de datos relacional fabricado por Oracle Corporation. Es una herramienta para la gestión de base de datos que adapta orientación a objetos en lo que respecta al tipos de datos
Arquitectura	Es la forma como se encuentra definido la estructura, funcionamiento e interacción de la base de datos	PostgreSQL es de arquitectura un modelo cliente/servidor y usa multiprocesos en vez de multihilos para garantizar la estabilidad del sistema. Es multiplataforma.	Oracle está basada en la arquitectura Cliente/Servidor para la gestión de Bases de Datos Relacionales. Además, soporta transacciones, es estable y multiplataforma.
Complejidad de sentencias sql (subconsultas)	Las consultas SQL complejas que usan subqueries	Postgres puede manejar rutinas complejas y reglas, consultas, subconsultas, vistas, soporta multiusuario, transacciones, optimizador de consultas, herencia y arreglos. No soporta selección de datos cruzados de diferentes tablas	Permite manejo de subconsultas, se implementó una utilidad que ayuda a simplificar las subconsultas complejas con el comando WITH. El comando WITH es una utilidad que materializa las subconsultas(Genera una tabla temporal)
Bloqueos y control de concurrencia	El control de concurrencia asegura que varios usuarios pueden acceder sobre el mismo dato al mismo tiempo	Control y bloqueo de concurrencia. Utiliza MVCC: Control de Concurrencia para Múltiples Versiones de PostgreSQL que se encarga de mantener copias sobre los datos de forma paralela para acelerar el sistema de escritura de datos a disco duro, haciendo un control de concurrencia entre las distintas versiones que se van escribiendo	Oracle maneja la concurrencia por medio de niveles de aislamiento (isolation levels) y MVCC

Tabla N° 39: Características de los gestores de base de datos (1)

Fuente: Investigadora

CARACTERISTICA	DESCRIPCION	POSTGRESQL	ORACLE
Integridad	Medidas necesarias para garantizar la validez de los datos almacenados.	PostgreSQL posee un gran control sobre la integridad referencial. Cumple totalmente con ACID (Atomicity, Consistency, Isolation, Durability). La integridad referencial se define con el uso combinado de claves primarias (primary keys) ó claves candidatas (candidate key) y clave foráneas (foreign key). Además permite empleo de transactions, rollbacks, foreignkeys ON DELETE CASCADE and ON UPDATE CASCADE.	Controla diferentes tipos de integridad: De entidad: cada fila en una tabla es única mediante claves primarias (PRIMARY KEY). De dominio: los datos cumplan con las especificaciones de dominio del diseño de la base de datos. Mediante exigencia de valor no nulo (NULL) y las restricciones (CONSTRAINT). De relaciones o referencial: garantiza que toda referencia tiene un destino. La forma para implementar ésta integridad en Oracle es mediante la restricción (foreignkey)
Encriptación	La encriptación es el proceso para volver ilegible para usuarios no autorizados la información. La información una vez encriptada sólo puede leerse aplicándole una clave.	Para encriptar los datos en la base de datos se puede usar un módulo adicional disponible en PostgreSQL que se llama pgcrypto. MD5, DES Y XDES	Tanto en su nivel básico como con el módulo <i>advanced security</i> que garantiza la seguridad. Usa algoritmos de encriptación como Algoritmos DES, Triple DES y RSA (el más rápido actualmente)
Estructuras de almacenamiento	Son las unidades lógicas de almacenamiento de una base de datos	PostgreSQL utiliza estructuras denominadas tablespaces que permiten definir las ubicaciones en el sistema de archivos de los objetos que representan las bases de datos	En Oracle los datos también se almacenan en tablespace, cada tablespace puede constar de 1 o más datafiles, es decir un tablespace puede almacenarse físicamente en uno o varios datafiles

Tabla N° 40: Características de los gestores de base de datos (2)

Fuente: Investigadora.

CARACTERISTICA	DESCRIPCION	POSTGRESQL	ORACLE
Almacenamiento de datos	Las técnicas empleadas para almacenar bases de datos son sumamente importantes para la velocidad de acceso y recuperación de datos.	Postgres tiene páginas desde 8kb hasta 32 kb y un registro no puede ser superior al tamaño de la página. Para datos grandes se comprime y parte en pequeños registros almacenados en tablas paralelas	Control Relacional de Bases de Datos. En dos niveles , fisico (datafiles) y lógico (tablespaces)
Particionamiento	consiste en subdividir un objeto de base de datos (tabla, índice, etc en partes más pequeñas teniendo cada parte un nombre y una forma de almacenamiento	PostgreSQL permite el particionamiento de tablas para mejorar el desempeño de la Base	Permite el uso de particiones para la mejora de la eficiencia, ciertas versiones admiten la administración de bases de datos distribuidas. Al instante de crear una tabla podemos elegir qué rangos de datos van a quedar almacenados en un tablespace u otro. Las ventajas que obtendremos son: Segmentos de datos más pequeños esto influye directamente en el rendimiento de las búsquedas porque cada partición es tratada como si fuera una tabla diferente; Índices más pequeños y respaldo más rápido
Herramientas gráficas de diseño y administración	Las herramientas de administración como las de diseño facilitan y agiliza la administración de estructuras y manipulación de datos	Existen algunas herramientas gráficas para administrar la base de datos PostgreSQL: pgAdmin ,pgAccess) y para hacer diseño de bases dedatos (Tora , Data Architect).	Oracle posee interfaces de administración de Base de datos como el EM (Oracle Enterprise Manager) y algunas herramientas de administración que no pertenecen a Oracle como TOAD.

Tabla N° 41: Características de los gestores de base de datos (3)

Fuente: Investigadora.

CARACTERISTICA	DESCRIPCION	POSTGRESQL	ORACLE
Recuperación de Base de datos (Recovery)	Las bases de datos tienen archivos de reconstrucción que les permite la recuperación de transacciones o estado completo en caso de fallo de base de datos.	PostgreSQL 8 usa Savepoints que permiten hacer un rolled back sin tener que repetir la transacción entera. Point in Time Recovery - Permite salvar el estado de la DB en momentos concretos	Archivos REDO que permite la recuperación de transacciones y permiten que Oracle coordine la forma de escritura en disco. Los backups completos de Base de datos pueden realizarse manualmente y por medio de la herramienta RMAN(RECOVERY MANAGER)
Seguridad	Un sistema gestor de base de datos cuenta con un subsistema de seguridad y autorización que se encarga de garantizar la seguridad de porciones de la BD contra el acceso no autorizado.	PostgreSQL cuenta con control de acceso mediante establecimiento de privilegios sobre roles y asignación de roles a usuarios.	Oracle ofrece varios niveles de seguridad: Seguridad en las cuentas para la validación de usuarios. Seguridad en el acceso para la base de datos Seguridad de nivel del sistema para gestionar privilegios globales.
Auditoría	Son registros que almacena la base de datos con información de acceso y ejecución de acciones sobre la base de datos.	La auditoría debe ser definida a nivel de tablas por el administrador por medio de triggers que almacenen registros de las acciones que se realizan.	Se puede auditar tres distintos tipos de acciones Intentos de inicio de sesión Accesos a objetos Acciones de base de datos

Tabla N° 42: Características de los gestores de base de datos (4)

Fuente: Investigadora.

CARACTERISTICA	DESCRIPCION	POSTGRESQL	ORACLE
Procesamiento de transacciones	Una transacción es una unidad lógica de ejecución de base de datos, es atómica e indivisible a pesar de estar compuesta por varios procesos es tratada como unidad.	Cumple con las reglas ACID. Durante las transacciones, existen restricciones diferidas (se verifican al final de la transacción) o inmediatas (se verifican en cada operación). Las transacciones permiten el paso entre dos estados consistentes manteniendo la integridad de los datos.	El control de transacciones de Oracle está diseñado para cumplir con las especificaciones ACID. Para garantizar la ejecución de transacciones realiza bloqueos exclusivos y bloqueos compartidos
Soporte de lenguaje	Especifica los tipos de caracteres que se pueda almacenar como datos.	PostgreSQL permite personalizar el soporte de lenguaje a cada usuario que se conecte en su respectivo local settings. Soporta juegos de caracteres internacionales, codificaciones de caracteres multibyte, Unicode y realiza ordenaciones dependiendo de la configuración de idioma local.	Permite cambiar y personalizar el idioma de base de datos haciendo uso de variables como NLS_CHARACTERSET,NLS_NCHAR_CHARACTERSET,NLS_LANGUAGE, NLS_TERRITORY,NLS_LANG
Escalabilidad	La escalabilidad es la propiedad deseable de un sistema, una red o un proceso, que indica su habilidad para extender el margen de operaciones sin perder calidad, o bien manejar el crecimiento continuo de trabajo de manera fluida, o bien para estar preparado para hacerse más grande sin perder calidad en los servicios ofrecidos	Es altamente escalable tanto en la cantidad bruta de datos que puede manejar como en el número de usuarios concurrentes que puede atender. Hay sistemas activos en producción con PostgreSQL que manejan más de 4 terabytes de datos	Oracle administra base de datos de hasta 100 TB sus características de estabilidad son muy efectivas.

Tabla N° 43: Características de los gestores de base de datos (5)

Fuente: Investigadora.

Los posibles puntos de fallo identificados fueron:

- Incompatibilidad en la migración en procedimientos o funciones donde se hace referencia a selección de datos cruzados de diferentes tablas.
- No existen paquetes en PostgreSQL.
- Los procedimientos como tal no existen en PostgreSQL, sino que son considerados como funciones que no devuelven valor.

Determinar herramientas de migración

Para poder seleccionar la herramienta se realizó pruebas a pequeña escala instalando las versiones de prueba de las herramientas propietarias pero que contenían las características de migración. Ver Anexo A4.

Luego de un análisis se estableció las siguientes ventajas de las herramientas de migración:

- Facilidad de uso
- Disponibilidad de soporte
- Costo de licenciamiento

Las tabla N° 44, tabla N° 45 y tabla N° 46 contienen la información de las ventajas y desventajas de las herramientas de migración.

HERRAMIENTA	VENTAJAS	DESVENTAJAS
SQLWays	Entorno gráfico	Costo de licencia privativa
	Contiene asistentes que permiten el uso intuitivo de la herramienta	En base de datos con estructuras complejas se reduce el porcentaje de automatización de migración
	Permite hacer una selección específica sobre los objetos de base de datos ya que cuenta con un formulario gráfico que brinda las facilidades de filtrado de objetos	
	Devuelve archivos por lotes para una ejecución de inserción sobre la base de datos	
	Ofrece servicios de migración y herramientas para Oracle, SQL server, MySQLPostgreSQL,IBM db2,Sybase, Informix y otras	
	SQLWays convierte procedimientos almacenados, funciones, disparadores, el esquema de base de datos (DDL), datos, vista	
	Usado para transferir gran volumen de información	
	Soporta diferentes sistemas gestores de base de datos	
	asegura casi el 100% un proceso automático de la migración	
	automatización para las tareas de migración DDL,	
	OPTIMIZACION DE CODIGO MIGRADO	
	Soporta diferentes tipos de base de datos incluido: Oracle, SQL Server, IBM DB2, MySQL, SchoonerSQL, Sybase, Informix, HP Neoview, Teradata	
	transforma variables de tipo varchar2 a character varyng	

Tabla N° 44: Herramientas de migración (1)

Fuente: Investigadora

HERRAMIENTA	VENTAJAS	DESVENTAJAS
Ora2Pg	Software libre (Redistribuible bajo GNU GPL)	Interfaz consola. Tipo texto
	Es multiplataforma	La salida es en scripts tipo texto que deben ser ejecutados sobre la base de datos.
	extrae todos los objetos de base de datos	Proporciona ayuda básica para convertir código PLSQL a PLPGSQL (aún necesita trabajo manual).
	Exporta el esquema completo de la base de datos (tablas, vistas, secuencias, índices), con clave única, principal, extranjera y comprueba las restricciones.	
	· Exporta subsidios y privilegios para los usuarios y grupos.	
	· Exporta una selección de tabla (especificando los nombres de tabla).	
	· Exporta funciones, desencadenadores, procedimientos, paquetes y cuerpos de paquete predefinidos.	
	Permite filtrado de datos para exportación con el uso de la cláusula where	
	· Admite objetos BLOB de Oracle como PG BYTEA.	
	· Exporta vistas de Oracle como tablas de PG.	
	· Exporta tipos Oracle definidos por el usuario.	

Tabla N° 45: Herramientas de migración (2)

Fuente: Investigadora

HERRAMIENTA	VENTAJAS	DESVENTAJAS
Enterprise Migration Toolkit	Licenciamiento software Propietario	costo de licenciamiento propietario
	Entorno gráfico	
	Permite migración online en la base de datos destino	
	Permite migración offline creando sentencias LDD en scripts	
	El costo relativo puede considerarse 3 veces inferior al licenciamiento Oracle	
	Permite Migración de aplicaciones de oracle a postgresql plus	
	Al ser parte de EnterpriseDB cuenta con las siguientes características:	
	<ul style="list-style-type: none"> • Estructuras de datos simples y complejas 	
	<ul style="list-style-type: none"> • Paquetes y funciones 	
	<ul style="list-style-type: none"> • Replicación Bidireccional 	
	<ul style="list-style-type: none"> • intuitivo de usar 	
	<ul style="list-style-type: none"> • se puede migrar datos o esquema , o las dos cosas al mismo tiempo 	
<ul style="list-style-type: none"> • Permite migración rápida 		

Tabla N° 46: Herramientas de migración (3)

Fuente: Investigadora

De las herramientas de migración revisadas se optó por usar Ora2Pg por su licencia de distribución libre y el soporte que tiene en la comunidad de usuarios PostgreSQL. Además de ser uno de los objetivos planteados por la empresa en el proyecto de migración el ahorro en costos y el deseo de incursionar en la línea de uso de software libre.

La extracción de estructuras realizada por Ora2Pg en archivos planos permite su modificación de sentencias DDL en caso de existir incompatibilidades de tipos.

La información empleada para descargar e instalar la herramienta Ora2Pg fue obtenida del sitio web de Ora2Pg administrado por Gilles Darold (información y dirección del sitio web adjunta en la bibliografía del presente proyecto)

La instalación detallada de la instalación y configuración de la herramienta de migración ora2Pg se encuentra en el Anexo A4.

Identificar información crítica

En base a la información recolectada en la etapa de levantamiento de requerimientos se determinó por nivel de acceso y tipo de transacción la criticidad de las tablas.

AEM3: Depuración y preparación de la base.

Se realizó la revisión de la información de igual manera basándonos en la criticidad e importancia de las tablas.

- Datos históricos almacenados.
- Información histórica de auditoría almacenada en la base de datos

6.7.2.2. PM2: Ejecución de la migración

AEJ1: Extracción de los datos

La extracción de los datos se realiza con la herramienta Ora2Pg, obteniendo la información por tipo de objetos, tablas, vistas, funciones y procedimientos.

Para utilizar Ora2Pg debemos cambiar la extensión del archivo ora2pg_pl a extensión “.pl”.

Se especificó los siguientes parámetros en el archivo de configuración ora2pg.conf como lo señala la tabla N° 37.

```
##### Ora2Pg Configuration file #####
# Set the Oracle home directory
ORACLE_HOME      D:\oracle\product\10.2.0\client_1

# Your SID should be declared on your tnsnames.ora file
ORACLE_DSN dbi:Oracle:host=192.168.20.100;sid=orcl
ORACLE_USER      origen
ORACLE_PWD origen

# Oracle schema/owner to use
SCHEMA           origen

# Type of export. Values can be the following keyword:
#   TABLE       Export tables
#   PACKAGE      Export packages
#   DATA        Export datas from table as INSERT
statement
(>= v8.4)
TYPE             TABLE

#By default all output is dump to STDOUT if not send
directly to postgresql
# database (see above). Give a filename to save export to
it. If you want
OUTPUT          output.sql
```

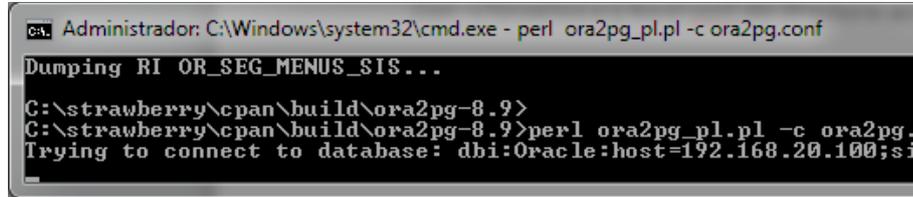
Tabla N° 47: archivo de configuración de Ora2PG

Fuente: Investigadora

- Se ejecuta Ora2Pg con el comando :

Perl ora2pg_pl.pl -c ora2pg.conf

Con `-c` forzamos a que lea el archivo.conf del directorio actual en el que se está trabajando, tal como lo demuestra el gráfico N° 30.



```
C:\Windows\system32\cmd.exe - perl ora2pg_pl.pl -c ora2pg.conf
Dumping RI OR_SEG_MENUS_SIS...
C:\strawberry\cpan\build\ora2pg-8.9>
C:\strawberry\cpan\build\ora2pg-8.9>perl ora2pg_pl.pl -c ora2pg.conf
Trying to connect to database: dbi:Oracle:host=192.168.20.100;si
```

Gráfico N° 30: Ejecución de Ora2Pg

Fuente: Investigadora

- Verificamos que el archivo de salida se haya creado. El gráfico N° 31 permite observar el archivo generado

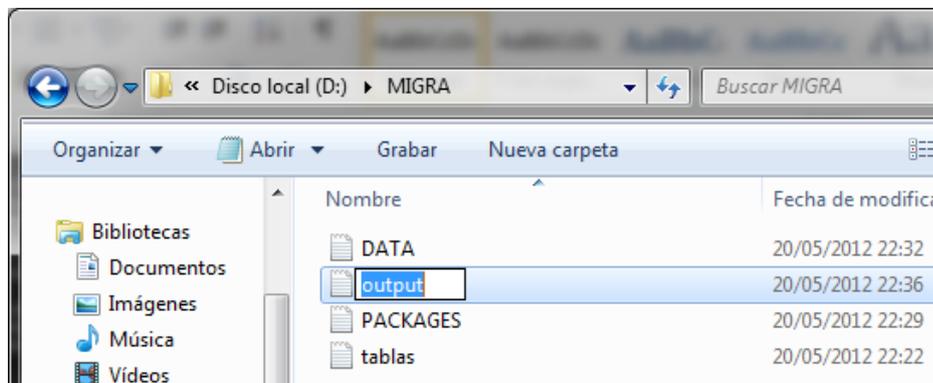


Gráfico N° 31: Verificación de creación de script

Fuente: Investigadora

- Se debe ejecutar Ora2Pg para extraer los diferentes tipos de objetos, ya que Ora2Pg solo obtiene un tipo de objeto a la vez. El gráfico N°32 nos demuestra la extracción de triggers de la base de datos.

```

C:\Windows\system32\cmd.exe
Dumping trigger OR_TGUP_BODEGAS...
Dumping trigger OR_TGUP_ENCABEZADOCOMPRAS1...
Dumping trigger OR_TGUP_ENCABEZADOCOMPRASTEMP1...
Dumping trigger OR_TGUP_ENCABEZADOREQ...
Dumping trigger OR_TGUP_ENCABEZADOTRANSF2...
Dumping trigger OR_TGUP_INSUMOS...
Dumping trigger OR_TGUP_KARDEX...
Dumping trigger OR_TGUP_OR_ARQUEO_CAJA1...
Dumping trigger OR_TGUP_PERIODOS...
Dumping trigger OR_TGUP_PROVEEDORES...
Dumping trigger OR_TGUP_PROUEE_SRI...
Dumping trigger OR_TGUP_RENGLONESCOMPRAS1...
Dumping trigger OR_TGUP_SECUENCIAS...

C:\strawberry\cpan\build\ora2pg-8.9>

```

Gráfico N° 32: Extracción de triggers

Fuente: Investigadora

La extracción de funciones de la base de datos a ser migrada se encuentra representada en el gráfico N° 33.

```

C:\Windows\system32\cmd.exe
Parsing function OR_FOBTIENE_BLOQUESRI_ACTIVO...
Dumping function OR_FOBTIENE_SECUENCIA...
Parsing function OR_FOBTIENE_SECUENCIA...
Dumping function OR_FOBTIENE_TURNO...
Parsing function OR_FOBTIENE_TURNO...
Dumping function OR_FPERIODO_ACTUAL...
Parsing function OR_FPERIODO_ACTUAL...
Dumping function OR_FPERMISO_BODEGA...
Parsing function OR_FPERMISO_BODEGA...
Dumping function OR_FUAL_CEDULA...
Parsing function OR_FUAL_CEDULA...
Dumping function OR_FUAL_RUC...
Parsing function OR_FUAL_RUC...

C:\strawberry\cpan\build\ora2pg-8.9>

```

Gráfico N° 33: Extracción de funciones

Fuente: Investigadora

El gráfico N° 34 describe la extracción de los procedimientos de la base de datos.

```

C:\Windows\system32\cmd.exe
Trying to connect to database: dbi:Oracle:host=192.168.20.100;sid=ora
Isolation level: SET TRANSACTION ISOLATION LEVEL SERIALIZABLE
Force Oracle to compile schema before code extraction
Retrieving procedures information...
Add procedures definition...
Dumping procedure OR_PACTUALIZA_SECUENCIA...
Parsing function OR_PACTUALIZA_SECUENCIA...
Dumping procedure OR_PBORRA_INSTANCIA...
Parsing function OR_PBORRA_INSTANCIA...
Dumping procedure OR_PCREA_INSTANCIA...
Parsing function OR_PCREA_INSTANCIA...
Dumping procedure OR_PINSERIA_STOCK_INSUMO...
Parsing function OR_PINSERIA_STOCK_INSUMO...

C:\strawberry\cpan\build\ora2pg-8.9>

```

Gráfico N° 34: Extracción de procedimientos

Fuente: Investigadora

- El gráfico N° 35 nos muestra los archivos obtenidos luego de usar la herramienta de migración.

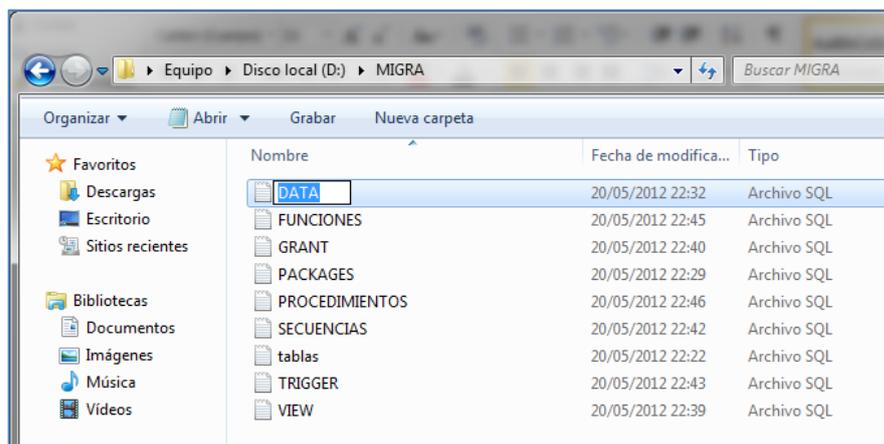


Gráfico N° 35: Archivos creados por ora2Pg

Fuente: Investigadora

AEJ2: Definición de estructuras destino:

Por medio del cliente de la base de datos PostgreSQL se realiza la definición de las tablas y más objetos de la base de datos a partir de los scripts obtenidos con la herramienta de migración. Para la definición de los objetos de base de datos usamos el cliente de PostgreSQL pgAdminIII. El gráfico N° 36 nos demuestra una imagen del cliente de PostgreSQL.

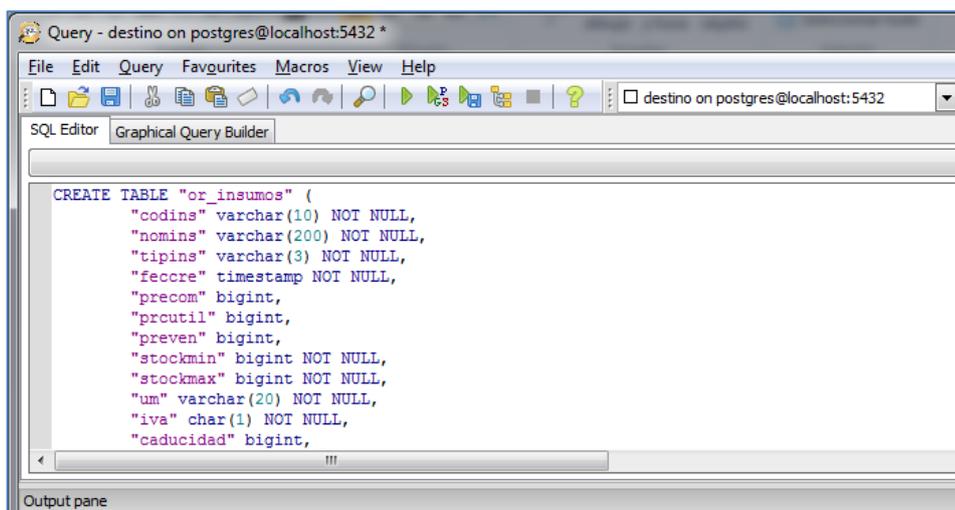


Gráfico N° 36: Cliente de PostgreSQL pgAdminIII

Fuente: Investigadora

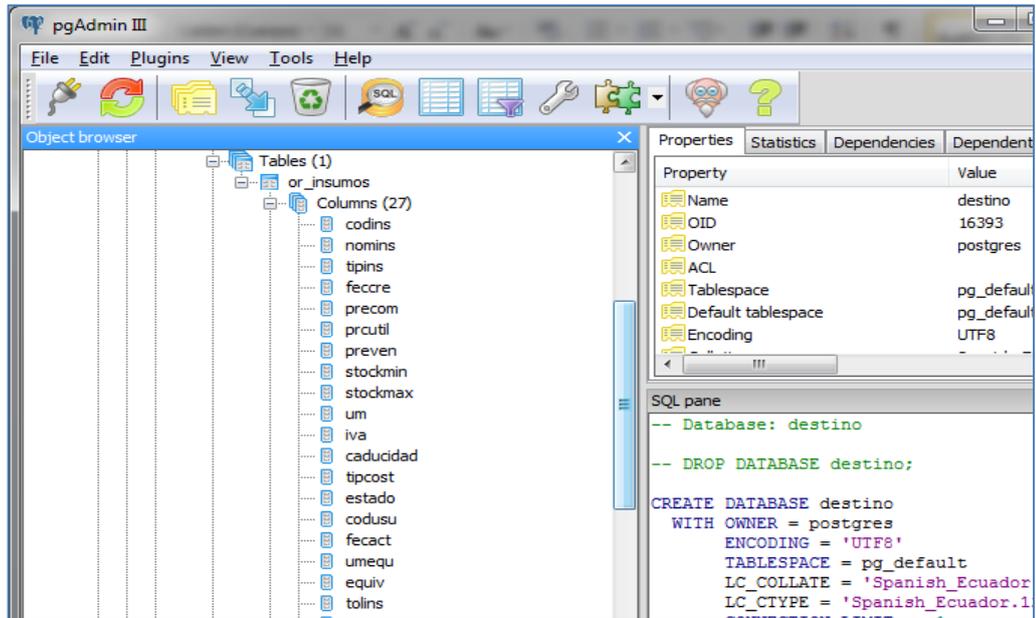


Gráfico N° 37: Verificación de creación de tablas

Fuente: Investigadora

Para observar los detalles técnicos y configuraciones ver el Manual Operativo de ejecución de migración (Anexo A5).

AEJ3: Depuración manual de la base de datos.

Al realizar la definición de objetos de la base de datos, comparamos tipos de datos de las tablas de origen y las tablas destino. Y realizamos las correcciones que fueron necesarias.

AEJ4: Planificación y pruebas de carga

Una vez que se determina que la estructura de la base de datos se encuentra lista se define las pruebas a realizar sobre la carga de datos. En éste caso la planificación realizada de pruebas está representado en la Tabla N° 38.

OBJETIVO	META	ACTIVIDADES	RESULTADOS	PLAZOS
Realizar pruebas unitarias de carga de datos en la base destino	Pruebas módulo Seguridad	Pruebas unitarias. Pruebas de integridad de datos	Hoja de registro de pruebas	3 dias
	Pruebas módulo Clientes	Pruebas unitarias. Pruebas de integridad de datos	Hoja de registro de pruebas	5 dias
	Pruebas módulo proveedores	Pruebas unitarias. Pruebas de integridad de datos	Hoja de registro de pruebas	3 dias
	Prueba modulo punto de venta	Pruebas unitarias. Pruebas de integridad de datos	Hoja de registro de pruebas	10 dias
	Prueba módulo Inventario	Pruebas unitarias. Pruebas de integridad de datos	Hoja de registro de pruebas	10 dias
	Análisis de registro de pruebas, detección y corrección de errores.	Clasificar errores. Corregir	LDD corregidas	12 dias
Realizar prueba de carga total de la base de datos	Realizar prueba inicial de carga total	Realizar cargas totales de la base de datos	Hoja de registro de pruebas	10 dias
	Prueba secundaria de carga total		Hoja de registro de pruebas	10 dias
	Otras validaciones		Hoja de registro de pruebas	10 dias
	Análisis de registro de pruebas, detección y corrección de errores.	Clasificar errores. Corregir.	LDD corregidas	12 dias

Tabla N° 48: Planificación de pruebas.

Fuente: Investigadora

AEJ5: Carga de la información

La carga de información fue satisfactoria al 95%.

6.7.2.3. PM3: Evaluación y control de calidad

La primera verificación se realizó a nivel de modelos conceptuales de base de datos origen y destino. Además se realizó una revisión de objetos de base de datos sobre cada gestor.

La información de tablas fue verificada programando sentencias de selección y sub consultas principalmente sobre la información crítica.

AC1: Análisis de integridad

La ejecución del análisis de integridad mediante inserciones, eliminaciones y actualizaciones manuales sobre las tablas de la base de datos destino con datos, se evaluó el cumplimiento de las mismas hasta llegar al 100% de cumplimiento de integridad

Se generó scripts con las operaciones para probar la integridad.

AC2: Análisis de consistencia

Se realizó las pruebas de caja negra sobre funcionalidad de base de datos y se depuró hasta cubrir el 100% de consistencia.

AC3: Evaluación de proceso global de migración

Una vez realizado el proceso de migración se evalúa el nivel de efectividad del proceso y de ser necesario se realiza afinamiento de la base de datos y depuración de información.

6.8. Administración

La aplicación y verificación de la metodología será administrada jerárquicamente por el siguiente personal de la empresa VSYSTEMS en el área de Sistemas.

El gráfico N° 28 representa como sería el proceso de administración de la propuesta.

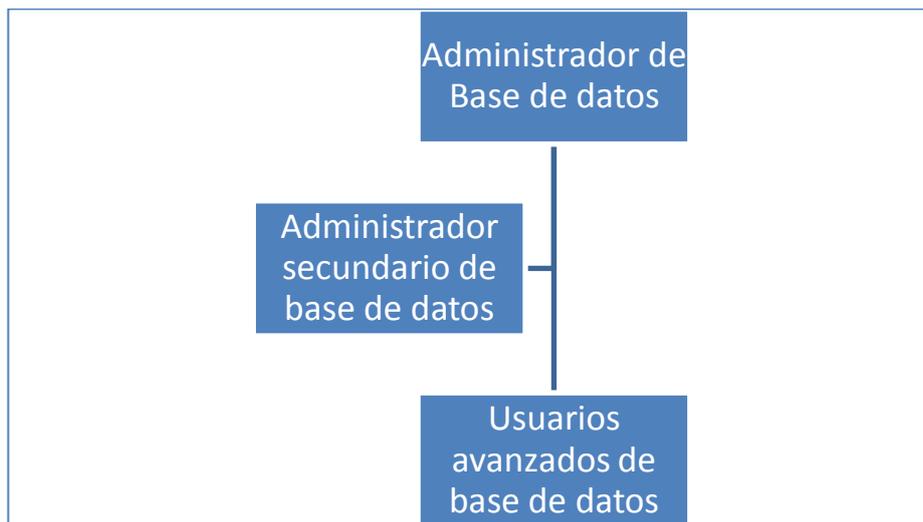


Gráfico N° 38: Administración

Fuente: Investigadora

6.9. Previsión de la evaluación

Corresponde a los directivos y al administrador de base de datos de la empresa VSYSTEMS realizar la aprobación y correcta aplicación de la metodología desarrollada en el presente proyecto.

7. CONCLUSIONES Y RECOMENDACIONES

7.1. Conclusiones

- El levantamiento de requerimientos dentro de un proceso de migración ayuda a determinar las pautas principales que guiarán el proceso y permite identificar objetivos empresariales planteados en el proceso migratorio.
- El uso de herramientas de migración facilita la ejecución del proceso mas no garantiza la automatización completa ya que pueden presentarse incompatibilidades en tipos de datos que deben ser resueltas de manera manual.
- La evaluación de control de calidad así como la aplicación correcta de pruebas de integridad asegura el cumplimiento de las reglas de integridad y consistencia de la información, dando como resultado una base de datos idónea y funcional.
- La existencia de documentación y un análisis de tablas que contienen información crítica facilita la evaluación del proceso de migración y la conservación de las reglas principales de lógica del negocio.
- La aplicación de una metodología desarrollada a medida permite ahorro de costos, ya que reduce tiempos improductivos de los sistemas asociados a la base de datos migrada.

7.2. Recomendaciones

- Para realizar un proceso de migración de base de datos es muy importante considerar todo el entorno del sistema relacionado a la misma, pues se debe identificar la interacción e importancia de cada tabla con una visión global y sistémica en la actividad empresarial.
- La realización de pruebas de carga permite detectar las incompatibilidades entre los tipos de datos originados luego del uso de la herramienta de migración, reduciendo así la tasa de errores y pronosticando posibles inconvenientes en el momento que se realiza la carga real de la base de datos.
- Las pruebas de integridad deben ser diseñadas y estructuradas cuidadosamente, cubriendo la mayor cantidad de posibles riesgos de manera que permitan lograr una buena calidad de datos.
- Se debe definir procesos de análisis de información e identificación de información crítica de base de datos para determinar políticas que permitan la preservación de la información y su consistencia en la empresa VSYSTEMS.
- Se debe realizar una revisión planificada y posterior limpieza de registros de auditoría para evitar almacenamiento de información innecesaria y reducir la cantidad de información a ser migrada. Además el proceso de depuración previa de datos y la correcta planificación permite el uso óptimo de los recursos.

MATERIALES DE REFERENCIA

BIBLIOGRAFÍA

- Achig, L. (2001). *Enfoques y Métodos de Investigación Científica*. Quito: AFEFCE.
- AnderEgg, E. (1995). *Técnicas de Investigación Social*. (24ª ed.). Argentina: El Cid Editor.
- Andino, P. (2005). *El Proyecto de Grado. Teoría y Gráficos*. Quito: P.H. Ediciones.
- Castaño, M. (2003). *Diseño de base de datos relacionales*. México: PrenticeHall.
- Castaño, M., Piattini, C y Velthuis, M. (2000). *Fundamentos y Modelos de Bases de Datos (2ª Ed.)* Editorial Ra-Ma,
- Elmasri, R. y Navathe, S (2002) *Fundamentos de Sistemas de Bases de Datos*. (3ª ed.) Addison-Wesley, Pearson Educación.
- Korth, H.; Silberschatz, A. (2006). *Fundamentos de base de datos (5ta Ed.)*. Madrid: McGraw-Hill.
- Kroenke, D. (2003). *Procesamiento de base de datos*. (8va Ed.) México: Prentice Hall.

Loney, K. y Theriault M. (2002). *Oracle 9i Manual del administrador* (2da Ed.). España: McGraw-Hill.

Ricardo, K. (2009). *Base de datos* (1ra Ed.). México: McGraw-Hill.

Abril, V. (s.f). *Técnicas e instrumentos de la Investigación*. Recuperado el 15 de Junio del 2011, de <http://vhabril.wikispaces.com/file/view/T%C3%A9cnicas+e+Instrumentos+de+la+Investigaci%C3%B3n.pdf>

Acronis (2010), *El costo oculto de la migración de datos*. Recuperado el 21 de Julio de 2011, de <http://www.liacolombia.com/pdf/costo-oculto-migracion-datos.pdf>

Anaya, V., Carsi, J., Pérez, J., Ramos, I. y Silva, J.(s.f). *Migración de datos automática a partir de la información de los esquemas conceptuales*. Recuperado el 11 de septiembre del 2011, de http://www.mhe.es/cf/ciclos_informatica/8448199421/archivos/unidad2_recurso2.pdf

Corona, S. (s.f). *Asegure una migración exitosa de base de datos*. Recuperado el 25 de abril del 2011, de <http://www.enterate.unam.mx/Articulos/2006/agosto/migracion.htm>

Darold, Gilles (s.f) *Ora2Pg, Migrates Oracle databases to PostgreSQL*. Recuperado el 20 de diciembre del 2011, de <http://ora2pg.darold.net/support.html>

Erickson, T. (2009, 3 de Noviembre). *Las herramientas de migración de datos llevan a las PYMEs al siguiente nivel: tutorial de migración y*

almacenamiento de datos por niveles. Recuperado el 2 de febrero del 2011, de <http://www.searchstorage.es/administracion-de-almacenamiento-de-datos/las-herramientas-de-migracion-de-datos-llevan-a-las-pymes-al-siguiente-nivel-tutorial-de-migracion-y-almacenamiento-de-datos-por-niveles/>

Hudicka, J (1998), *An overview of data Migration Methodology*, Recuperado el 14 de abril del 2012, desde http://www.dulcian.com/articles/overview_data_migration_methodology.htm

Subsecretaria de Informática de la República del Ecuador (s.f). *Estrategia de Migración a Software Libre*. Recuperado el 15 de enero del 2011, de <http://www.informatica.gob.ec/index.php/software-libre/estrategia-de-migracion>

Primer Taller Regional de capacitación PROVAC (2008), *Importancia del Proceso Conservar de Migración de la Base de datos*. Recuperado el 12 de febrero del 2011 de http://www.censia.salud.gob.mx/descargas/Estrategias/4_importancia_de_la_migracion.pdf

Subsecretaria de Informática del Ministerio de Telecomunicaciones e Información del Ecuador (s.f). *Alternativas de Software Libre para Software Privativo*. Recuperado el 15 de enero del 2011, de <http://www.informatica.gob.ec/index.php/software-libre/alternativas>

pgFoundry (s.f) the PostgreSQL Development Group's site. *Ora2Pg*. Recuperado el 18 de diciembre del 2011, de <http://pgfoundry.org/>

Quiñones, E (s.f). *Sistemas de base de datos libres para el entorno empresarial*. Recuperado el 10 de marzo del 2011, de http://www.eqsoft.net/presentas/sistemas_administradores_de_bases_de_datos_libres_para_el_entorno_

empresarial.pdf

Quiñones, E (s.f). *Consideraciones Técnicas para un proceso de Migración*. Recuperado el 30 de agosto del 2011, de http://www.eqsoft.net/presentas/consideraciones_tecnicas_para_una_migracion_a_software_libre.pdf

Smith, G., Treat, R., & Browne, C. (09 de 05 de 2012). *Tuning Your PostgreSQL Server/es*. Recuperado el 18 de junio de 2012, de http://wiki.postgresql.org/wiki/Tuning_Your_PostgreSQL_Server/es

Singh, Gurjeet (2011-05-19), *Oracle to Postgres Migration Considerations, Hurdles, and possible Solutions*. Recuperado en 08 de enero del 2012, desde http://www.pgcon.org/2011/schedule/attachments/205_Oracle_to_Postgres_Migration.pdf

Wikipedia (2010, 29 de septiembre), *Técnicas de Migración de datos*. Recuperado el 10 de junio del 2011, de http://es.wikipedia.org/wiki/Migraci%C3%B3n_de_datos

ANEXOS

ANEXO A1: Especificación de *Requisitos* según el estándar de *IEEE* 830 (Instituto de Ingenieros Eléctricos y Electrónicos) 22 de octubre del 2008. Obtenida desde la dirección electrónica <http://www.fdi.ucm.es/profesor/gmendez/docs/is0809/ieee830.pdf>

Especificación de Requisitos según el estándar de IEEE 830

IEEE Std. 830-1998

22 de Octubre de 2008

Resumen

Este documento presenta, en castellano, el formato de Especificación de Requisitos Software (ERS) según la última versión del estándar IEEE 830. Según IEEE, un buen Documento de Requisitos, pese a no ser obligatorio que siga estrictamente la organización y el formato dados en el estándar 830, sí deberá incluir, de una forma o de otra, toda la información presentada en dicho estándar. El estándar de IEEE 830 no está libre de defectos ni de prejuicios, y por ello ha sido justamente criticado por múltiples autores y desde múltiples puntos de vista, llegándose a cuestionar incluso si es realmente un estándar en el sentido habitual que tiene el término en otras ingenierías. El presente documento no pretende pronunciarse ni a favor ni en contra de unos u otros: tan sólo reproduce, con propósitos fundamentalmente docentes, cómo se organizaría un Documento de Requisitos según el estándar IEEE 830.

Índice

1. Introducción	3
1.1. Propósito	3
1.2. Ámbito del Sistema	3
1.3. Definiciones, Acrónimos y Abreviaturas	3
1.4. Referencias	3
1.5. Visión General del Documento	4
2. Descripción General	4
2.1. Perspectiva del Producto	4
2.2. Funciones del Producto	4
2.3. Características de los Usuarios	5
2.4. Restricciones	5
2.5. Suposiciones y Dependencias	5
2.6. Requisitos Futuros	6
3. Requisitos Específicos	6
3.1. Interfaces Externas	7
3.2. Funciones	7
3.3. Requisitos de Rendimiento	9
3.4. Restricciones de Diseño	9
3.5. Atributos del Sistema	9
3.6. Otros Requisitos	9

1. Introducción

En esta sección se proporcionará una introducción a todo el documento de Especificación de Requisitos Software(ERS). Consta de varias subsecciones: propósito, ambito del sistema, definiciones, referencias y visión general del documento.

1.1. Propósito

En esta subsección se definirá el propósito del documento ERS y se especificará a quién va dirigido el documento.

1.2. Ámbito del Sistema

En esta subsección:

- Se podrá dar un nombre al futuro sistema (p.ej. MiSistema)
- Se explicará lo que el sistema hará y lo que no hará.
- Se describirán los beneficios, objetivos y metas que se espera alcanzar con el futuro sistema.
- Se referenciarán todos aquellos documentos de nivel superior (p.e. en Ingeniería de Sistemas, que incluyen Hardware y Software, debería mantenerse la consistencia con el documento de especificación de requisitos globales del sistema, si existe).

1.3. Definiciones, Acrónimos y Abreviaturas

En esta subsección se definirán todos los términos, acrónimos y abreviaturas utilizadas en la ERS.

1.4. Referencias

En esta subsección se mostrará una lista completa de todos los documentos referenciados en la ERS.

1.5. Visión General del Documento

Esta subsección describe brevemente los contenidos y la organización del resto de la ERS.

2. Descripción General

En esta sección se describen todos aquellos factores que afectan al producto y a sus requisitos. No se describen los requisitos, sino su contexto. Esto permitirá definir con detalle los requisitos en la sección 3, haciendo que sean más fáciles de entender.

Normalmente, esta sección consta de las siguientes subsecciones: Perspectiva del producto, funciones del producto, características de los usuarios, restricciones, factores que se asumen y futuros requisitos.

2.1. Perspectiva del Producto

Esta subsección debe relacionar el futuro sistema (producto software) con otros productos. Si el producto es totalmente independiente de otros productos, también debe especificarse aquí. Si la ERS define un producto que es parte de un sistema mayor, esta subsección relacionará los requisitos del sistema mayor con la funcionalidad del producto descrito en la ERS, y se identificarán las interfaces entre el producto mayor y el producto aquí descrito. Se recomienda utilizar diagramas de bloques.

2.2. Funciones del Producto

En esta subsección de la ERS se mostrará un resumen, a grandes rasgos, de las funciones del futuro sistema. Por ejemplo, en una ERS para un programa de contabilidad, esta subsección mostrará que el sistema soportará el mantenimiento de cuentas, mostrará el estado de las cuentas y facilitará la facturación, sin mencionar el enorme detalle que cada una de estas funciones requiere.

2.3. Características de los Usuarios

Esta subsección describirá las características generales de los usuarios del producto, incluyendo nivel educacional, experiencia y experiencia técnica.

2.4. Restricciones

Esta subsección describirá aquellas limitaciones que se imponen sobre los desarrolladores del producto

- Políticas de la empresa
- Limitaciones del hardware
- Interfaces con otras aplicaciones
- Operaciones paralelas
- Funciones de auditoría
- Funciones de control
- Lenguaje(s) de programación
- Protocolos de comunicación
- Requisitos de habilidad
- Criticalidad de la aplicación
- Consideraciones acerca de la seguridad

2.5. Suposiciones y Dependencias

Esta subsección de la ERS describirá aquellos factores que, si cambian, pueden afectar a los requisitos. Por ejemplo, los requisitos pueden presuponer una cierta organización de ciertas unidades de la empresa, o pueden presuponer que el sistema correrá sobre cierto sistema operativo. Si cambian dichos detalles en la organización de la empresa, o si cambian ciertos detalles técnicos, como el sistema operativo, puede ser necesario revisar y cambiar los requisitos.

3. Requisitos Específicos

Esta sección contiene los requisitos a un nivel de detalle suficiente como para permitir a los diseñadores diseñar un sistema que satisfaga estos requisitos, y que permita al equipo de pruebas planificar y realizar las pruebas que demuestren si el sistema satisface, o no, los requisitos. Todo requisito aquí especificado describirá comportamientos externos del sistema, perceptibles por parte de los usuarios, operadores y otros sistemas. Esta es la sección más larga e importante de la ERS. Deberán aplicarse los siguientes principios:

- El documento debería ser perfectamente legible por personas de muy distintas formaciones e intereses.
- Deberán referenciarse aquellos documentos relevantes que poseen alguna influencia sobre los requisitos.
- Todo requisito deberá ser unívocamente identificable mediante algún código o sistema de numeración adecuado.
- Lo ideal, aunque en la práctica no siempre realizable, es que los requisitos posean las siguientes características:
 - Corrección: La ERS es correcta si y sólo si todo requisito que figura aquí (y que será implementado en el sistema) refleja alguna necesidad real. La corrección de la ERS implica que el sistema implementado será el sistema deseado.
 - No ambiguos: Cada requisito tiene una sola interpretación. Para eliminar la ambigüedad inherente a los requisitos expresados en lenguaje natural, se deberán utilizar gráficos o notaciones formales. En el caso de utilizar términos que, habitualmente, poseen más de una interpretación, se definirán con precisión en el glosario.
 - Completos: Todos los requisitos relevantes han sido incluidos en la ERS. Conviene incluir todas las posibles respuestas del sistema a los datos de entrada, tanto válidos como no válidos.

- **Consistentes:** Los requisitos no pueden ser contradictorios. Un conjunto de requisitos contradictorio no es implementable.
- **Clasificados:** Normalmente, no todos los requisitos son igual de importantes. Los requisitos pueden clasificarse por importancia (esenciales, condicionales u opcionales) o por estabilidad (cambios que se espera que afecten al requisito). Esto sirve, ante todo, para no emplear excesivos recursos en implementar requisitos no esenciales.
- **Verificables:** La ERS es verificable si y sólo si todos sus requisitos son verificables. Un requisito es verificable (testeable) si existe un proceso finito y no costoso para demostrar que el sistema cumple con el requisito. Un requisito ambiguo no es, en general, verificable. Expresiones como a veces, bien, adecuado, etc. introducen ambigüedad en los requisitos. Requisitos como “en caso de accidente la nube tóxica no se extenderá más allá de 25Km” no es verificable por el alto costo que conlleva.
- **Modificables:** La ERS es modificable si y sólo si se encuentra estructurada de forma que los cambios a los requisitos pueden realizarse de forma fácil, completa y consistente. La utilización de herramientas automáticas de gestión de requisitos (por ejemplo RequisitePro o Doors) facilitan enormemente esta tarea.
- **Trazables:** La ERS es trazable si se conoce el origen de cada requisito y se facilita la referencia de cada requisito a los componentes del diseño y de la implementación. La trazabilidad hacia atrás indica el origen (documento, persona, etc.) de cada requisito. La trazabilidad hacia delante de un requisito R indica qué componentes del sistema son los que realizan el requisito R.

3.1. Interfaces Externas

Se describirán los requisitos que afecten a la interfaz de usuario, interfaz con otros sistemas (hardware y software) e interfaces de comunicaciones.

3.2. Funciones

Esta subsección (quizá la más larga del documento) deberá especificar todas aquellas acciones (funciones) que deberá llevar a cabo el software. Nor-

malmente (aunque no siempre), son aquellas acciones expresables como “el sistema deberá ...”. Si se considera necesario, podrán utilizarse notaciones gráficas y tablas, pero siempre supeditadas al lenguaje natural, y no al revés.

Es importante tener en cuenta que, en 1983, el Estándar de IEEE 830 establecía que las funciones deberían expresarse como una jerarquía funcional (en paralelo con los DFDs propuestos por el análisis estructurado). Pero el Estándar de IEEE 830, en sus últimas versiones, ya permite organizar esta subsección de múltiples formas, y sugiere, entre otras, las siguientes:

- Por tipos de usuario: Distintos usuarios poseen distintos requisitos. Para cada clase de usuario que exista en la organización, se especificarán los requisitos funcionales que le afecten o tengan mayor relación con sus tareas.
- Por objetos: Los objetos son entidades del mundo real que serán reflejadas en el sistema. Para cada objeto, se detallarán sus atributos y sus funciones. Los objetos pueden agruparse en clases. Esta organización de la ERS no quiere decir que el diseño del sistema siga el paradigma de Orientación a Objetos.
- Por objetivos: Un objetivo es un servicio que se desea que ofrezca el sistema y que requiere una determinada entrada para obtener su resultado. Para cada objetivo o subobjetivo que se persiga con el sistema, se detallarán las funciones que permitan llevarlo a cabo.
- Por estímulos: Se especificarán los posibles estímulos que recibe el sistema y las funciones relacionadas con dicho estímulo.
- Por jerarquía funcional: Si ninguna de las anteriores alternativas resulta de ayuda, la funcionalidad del sistema se especificará como una jerarquía de funciones que comparten entradas, salidas o datos internos. Se detallarán las funciones (entrada, proceso, salida) y las subfunciones del sistema. Esto no implica que el diseño del sistema deba realizarse según el paradigma de Diseño Estructurado.

Para organizar esta subsección de la ERS se elegirá alguna de las anteriores alternativas, o incluso alguna otra que se considere más conveniente. Deberá, eso sí, justificarse el porqué de tal elección.

3.3. Requisitos de Rendimiento

Se detallarán los requisitos relacionados con la carga que se espera tenga que soportar el sistema. Por ejemplo, el número de terminales, el número esperado de usuarios simultáneamente conectados, número de transacciones por segundo que deberá soportar el sistema, etc.

También, si es necesario, se especificarán los requisitos de datos, es decir, aquellos requisitos que afecten a la información que se guardará en la base de datos. Por ejemplo, la frecuencia de uso, las capacidades de acceso y la cantidad de registros que se espera almacenar (decenas, cientos, miles o millones).

3.4. Restricciones de Diseño

Todo aquello que restrinja las decisiones relativas al diseño de la aplicación: Restricciones de otros estándares, limitaciones del hardware, etc.

3.5. Atributos del Sistema

Se detallarán los atributos de calidad (las “ilities”) del sistema: Fiabilidad, mantenibilidad, portabilidad, y, muy importante, la seguridad. Deberá especificarse qué tipos de usuario están autorizados, o no, a realizar ciertas tareas, y cómo se implementarán los mecanismos de seguridad (por ejemplo, por medio de un *login* y una *password*).

3.6. Otros Requisitos

Cualquier otro requisito que no encaje en otra sección.

4. Apéndices

Pueden contener todo tipo de información relevante para la ERS pero que, propiamente, no forme parte de la ERS. Por ejemplo:

1. Formatos de entrada/salida de datos, por pantalla o en listados.
2. Resultados de análisis de costes.
3. Restricciones acerca del lenguaje de programación.

ANEXO A2: Documento de requerimientos del proceso de migración de datos de la empresa VSYSTEMS

Documento de requerimientos del proceso de migración de datos de la empresa VSYSTEMS

El presente documento de requerimientos para migración de la empresa VSYSTEMS está basado en la estructura propuesta por el IEEE para documentos de requerimientos.

1. Introducción

La empresa VSYSTEMS de la provincia de Santo Domingo de los Tsáchilas se dedica a desarrollo y venta de software, además de ofrecer consultoría en sistemas informáticos. Esta empresa cuenta con algunas aplicaciones, entre ellas se encuentra el sistema de puntos de venta e inventario que trabaja con una base de datos Oracle10g. Este sistema se encuentra instalado en varias empresas clientes y la base de datos respectiva es administrada por la empresa VSYSTEMS dentro de cada empresa cliente.

El proceso de migración que se va a emprender tiene como origen el gestor de Base de datos Oracle10g y como destino PostgreSQL.

1.1 Propósito del documento de requerimientos

El presente documento fue realizado para determinar los requerimientos del proceso de migración de base de datos que será realizado en la empresa VSYSTEMS.

1.2 Alcance

El presente documento especifica la migración de base de datos de propiedad de la empresa VSYSTEMS correspondiente al sistema de inventarios y punto de venta desde un gestor Oracle 10g a un gestor PostgreSQL.

1.3 Definiciones, acrónimos y abreviaturas

Base de datos: Se conoce como base de datos a un conjunto de datos que representan información relacionada a un mismo contexto almacenado de manera sistemática.

Migración de datos: Proceso que consiste en trasladar información desde un gestor de base de datos a otro gestor de diferente versión o tipo.

Integridad de datos: Se refiere a que las características de los datos deben ser correctas para que la información se mantenga relacionada es decir la información debe ser exacta, coherente y válida.

PK: Llave primaria de una tabla

ER: Modelo entidad relación

Gestor de base de datos (SGBD): Un gestor de base de datos está constituido por aplicaciones cuyo objetivo es proporcionar una interfaz entre la base de datos, el usuario y las aplicaciones. Dentro de un gestor se utiliza lenguaje de definición de datos y lenguajes de manipulación de datos.

Oracle 10G: La base de datos Oracle es un sistema de administración de base de datos relacionales (RDBMS). El modelo relacional consiste en utilizar tablas bidimensionales para almacenar la información. Consta de tres elementos básicos: tablas, conjunto de operadores para manipular esas tablas y reglas de integridad

PostgreSQL: es un sistema de gestión de base de datos relacional orientada a objetos y libre, publicado bajo la licencia BSD. El desarrollo de PostgreSQL no es manejado por una empresa y/o persona, sino que es dirigido por una comunidad de desarrolladores que trabajan de forma desinteresada, altruista, libre y/o apoyados por organizaciones comerciales. Dicha comunidad es denominada el PGDG (*PostgreSQL Global DevelopmentGroup*).

1.4 Referencias

El presente documento contiene además las características generales y específicas que se realizarán en el proceso de migración.

2. DESCRIPCION GENERAL

2.1 Perspectiva de la migración

La migración será realizada entre gestores Oracle 10 g y PostgreSQL asegurando la calidad del proceso y obteniendo finalmente una base de datos funcional, segura y de buen rendimiento cuya información sea válida y exacta.

2.2 Funciones de la migración

La migración de Base de datos comprenderá la estructura de base de datos correspondiente a:

- Tablas
- Procedimientos

- Índices
- Usuarios
- Triggers
- Funciones
- Reglas
- Procedimientos de respaldo, logs y seguridad.

Además de la estructura se migrará el contenido de las respectivas tablas debidamente relacionado.

2.3 Características del usuario

La base de datos específica a migrar de la empresa VSYSTEMS interactúa con el sistema de punto de venta y control de inventarios, ejecutándose sobre plataforma Windows y con un gestor de Base de datos Oracle 10g

En base a encuestas y entrevistas realizadas podemos conocer que el proceso migratorio en la empresa se origina por obtener ahorro en costos de licenciamiento y por la búsqueda de innovación tecnológica.

El crecimiento anual estimado de la base de datos es del 20%.

2.4 Restricciones generales

- Se asegura la total confidencialidad sobre la información que intervendrá en el presente proceso de migración de datos.
- Truncamiento de datos y/o pérdida de información

2.5 Suposiciones y dependencias

- Las funcionalidades que no sean migradas automáticamente por la herramienta de migración deberán ser reprogramadas en el Nuevo gestor.
- Las tablas que por incompatibilidad de tipo de datos no sean migradas automáticamente deben ser estructuradas y creadas en el Nuevo gestor.

3. REQUERIMIENTOS ESPECIFICOS

3.1 Requerimientos de seguridad

- Que se migre la información de usuarios y roles respectivamente.
- Se debe conservar las características de seguridad y control de usuarios.
- Conservar los niveles de seguridad o mejorarlos con respecto a la base de datos origen.

3.2 Requerimientos de almacenamiento

- De ser posible se mantendrá la misma estructura lógica.
- Que el gestor de la base de datos resultante se encuentre operativo sobre plataforma Windows.

3.3 Requerimientos de funcionalidad

- Que se conserve la consistencia de la información
- Que se conserve la integridad de la información
- Que la base de datos resultante esté en un estado funcional óptimo
- Que se conserve las reglas de negocio aplicadas
- Que no se pierda la numeración de las secuencias automáticas que poseen las tablas.
- Las reglas y restricciones de los datos deben estar presentes en cada tabla migrada.

ANEXO A3: Respaldo de configuraciones y datos de PostgreSQL

Respaldo de configuraciones

Se debe respaldar los archivos de sistema operativo correspondiente a la configuración de PostgreSQL. Los archivos a respaldar son:

- `postgresql.conf` que contiene los parámetros de configuración básicos de la base de datos.
- `pg_hba.conf` que registra el acceso y restricciones a usuarios o direcciones IP.

Respaldo de datos

En PostgreSQL existen dos maneras de obtener respaldos a nivel de información de Base de datos:

- Copiando los archivos donde se almacenan los datos desde el sistema operativo, en este caso sería los archivos ubicados en

```
C:\Program Files\PostgreSQL\9.0\data
```

Para este tipo de respaldo se debe detener la base

- Por medio de la generación de un script que permita recrear toda la base de datos, mediante un programa cliente propio de PostgreSQL `pg_dump` que puede ser ejecutado por el administrador de la base de datos, para la restauración podemos usar la sentencia `pg_restore`

Para la extracción del respaldo

```
pg_dump destino>
```

Para restaurar un respaldo

```
$ pg_restore -d dest_nuevo respaldo1.sql
```

En donde `-d` hace referencia que se debe cargar los datos a una base de datos específica.

ANEXO A4: Instalación y pruebas de herramientas de migración

- Instalación y uso de SQLWays

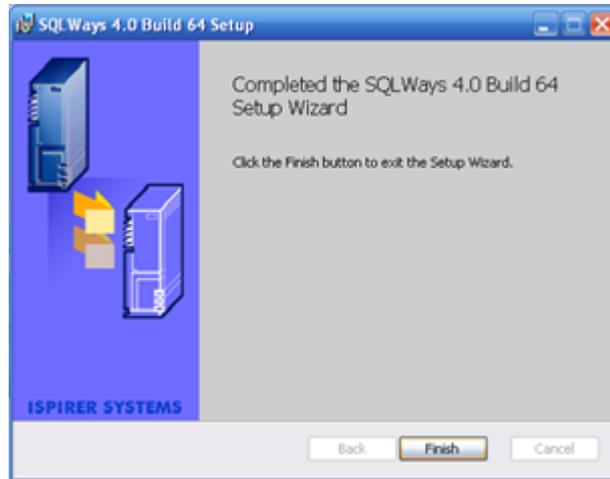


Gráfico N° 39: Pantalla inicial de instalación herramienta SQLWays

Fuente: Investigadora

El uso de la herramienta para migración es el siguiente:

- Se configura la conexión a la base de datos origen y a la base de datos destino.

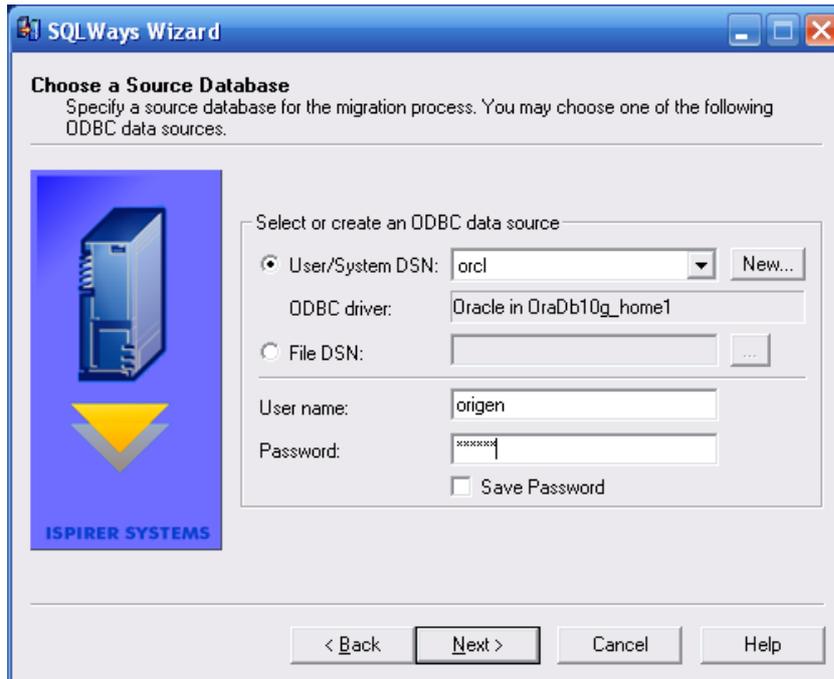


Gráfico N° 40: Pantalla de configuración del origen en SQLWays

Fuente: Investigadora

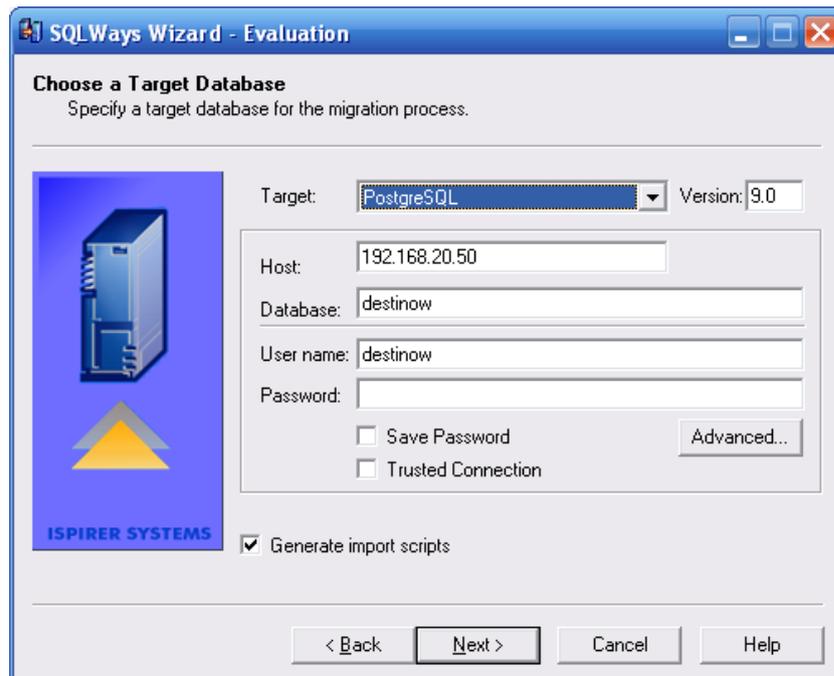


Gráfico N° 41: Pantalla de selección del destino de migración de SQLWays

Fuente: Investigadora

- Se despliega un formulario que permite seleccionar los objetos a ser migrados:

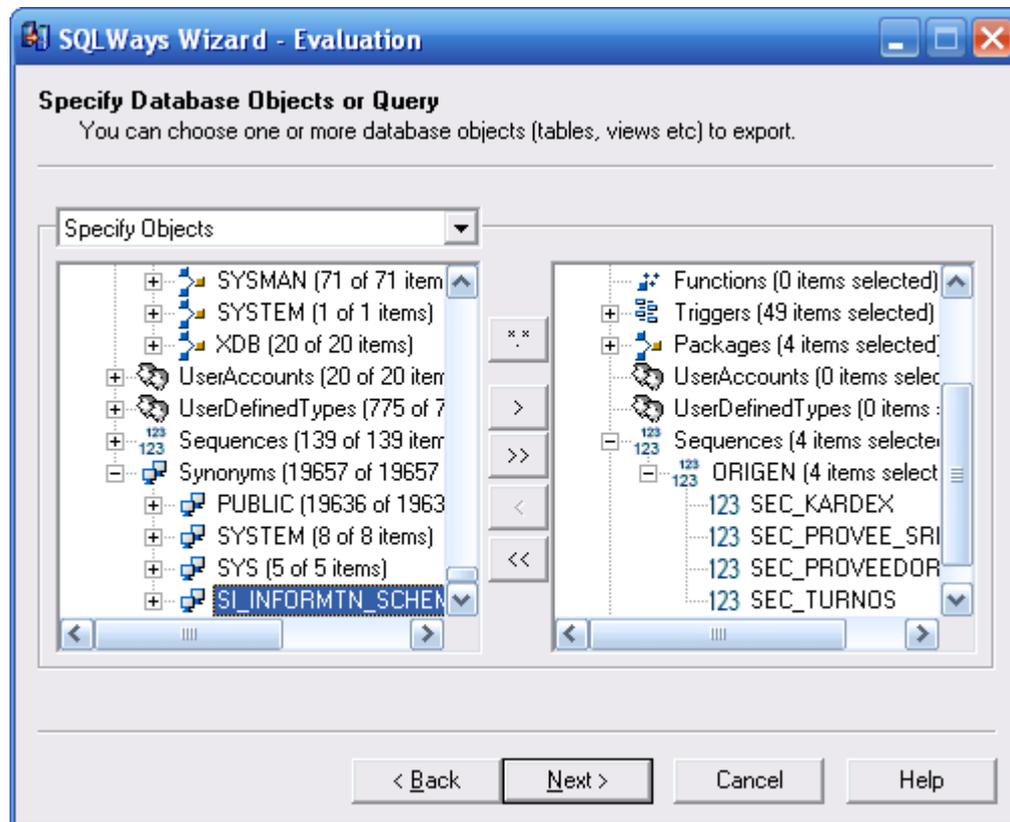


Gráfico N° 42: Pantalla de selección de objetos de migración de SQLWays

Fuente: Investigadora

- Especificamos la ruta donde se almacenarán los scripts y archivos de ejecución por lotes y luego se realiza la revisión de información y datos a migrarse

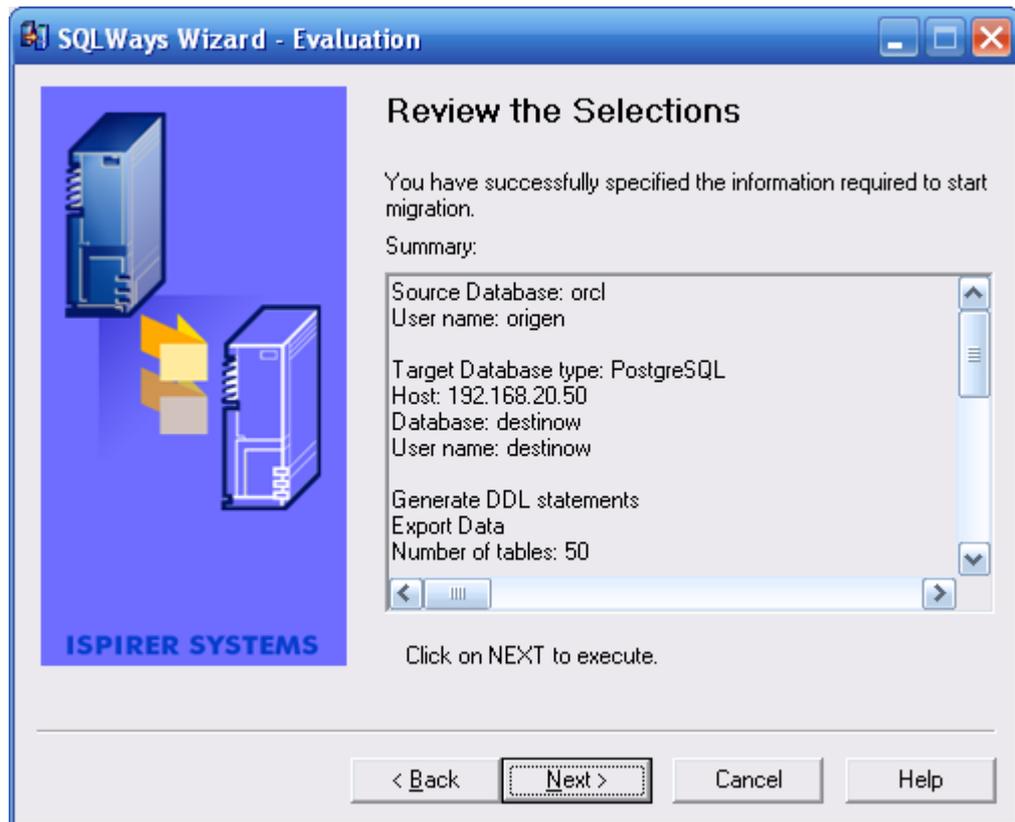


Gráfico N° 43: Pantalla de instalación de SQLWays

Fuente: Investigadora

- Finalmente los archivos de salida son almacenados en el directorio especificado para ese usuario. Se deben ejecutar los archivos.bat para la inserción de datos en la base de datos destino.

Instalación Enterprise DB

Obtuvimos el instalador de un demo de Enterprise DB, como un archivo ejecutable para ser instalado bajo sistema operativo Windows

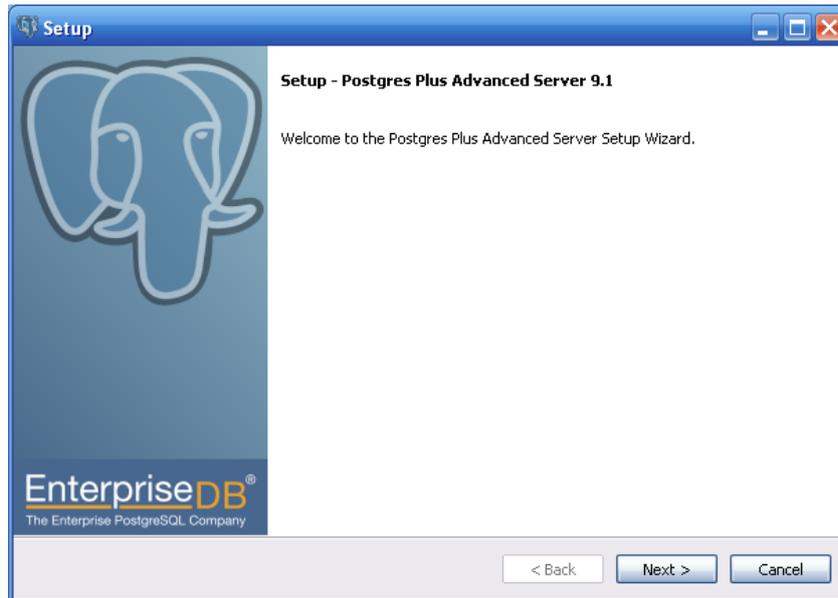


Gráfico N° 44: Pantalla de instalación de Postgres Plus

Fuente: Investigadora

En la siguiente pantalla se muestra los componentes disponibles para la instalación de EnterpriseDB.

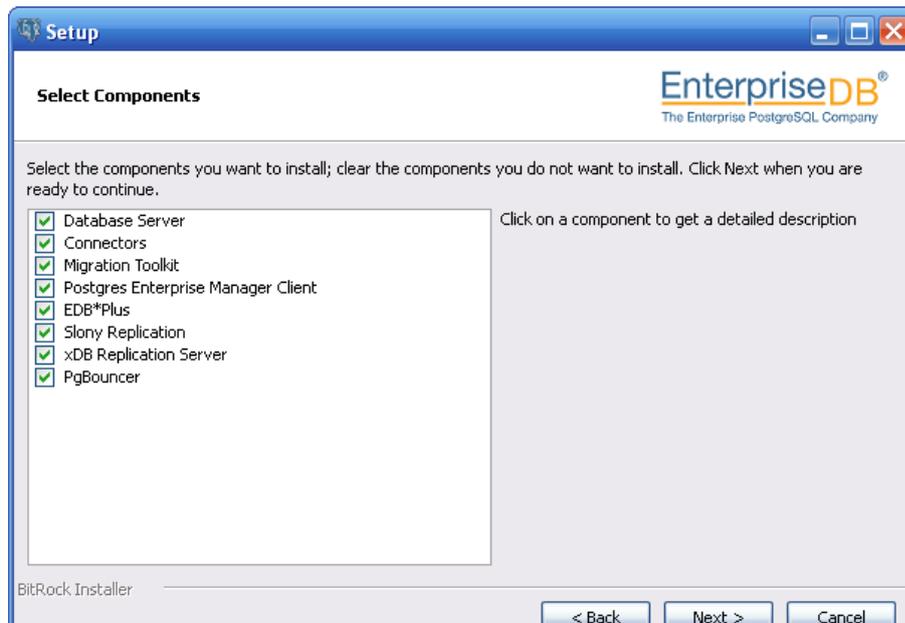


Gráfico N° 45: Pantalla de instalación de componentes EnterpriseDB

Fuente: Investigadora

El instalador nos permite configurar el servidor para optimizar el funcionamiento en determinada carga de trabajo. Además permite la configuración avanzada del servidor.

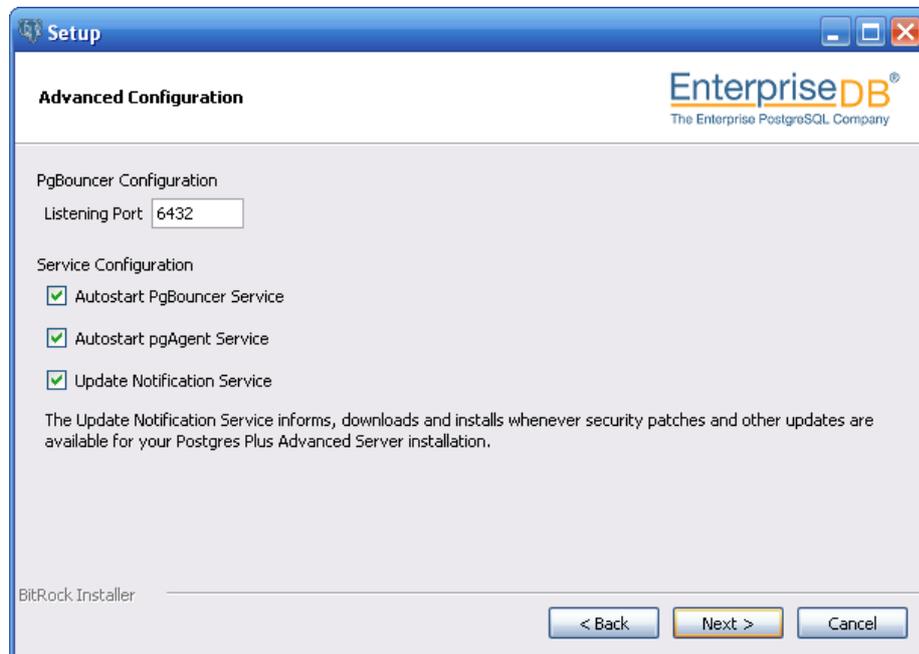


Gráfico N° 46: Pantalla de instalación de configuración puerto Enterprise DB

Fuente: Investigadora

ANEXO A5: Manual operativo ejecución de la migración

Requisitos previos

Hardware

- Instalación de red
- Servidor de base de datos origen
- PC cliente con cliente postgresSQL

Software

- Base de datos Oracle instalada en sistema operativo Windows

- Pc para migración
- Cliente de Oracle para la versión 10.2.0.1
 - Instalación de StrawberryPerl
 - Instalación y configuración de Ora2Pg
 - Tener conocimiento de administración de gestor Oacle y Postgresql.

1. Configuración de variables de entorno

Para la correcta instalación y funcionamiento de los módulos de Perl y uso de Ora2Pg se debe establecer variables de entorno luego de instalar el cliente de Oracle y la Base de datos de PostgreSQL. La definición de variables de entorno fueron:

- Variables de entorno para conexión Oracle

```
ORACLE_HOME=D:\oracle\product\10.2.0\client_1
ORACLE_SID=ORCL
ORACLE_USERID=origen/origen
ORACLE_DSN= dbi:Oracle:host=192.168.20.100;sid=orcl
TNS_ADMIN= D:\oracle\product\10.2.0\client_1\NETWORK\ADMIN
LD_LIBRARY_PATH=D:\oracle\product\10.2.0\client_1\LIB
NLS_LANG==.WE8ISO8859P15
En la variable PATH se debe añadir el directorio bin del Cliente
Oracle, en éste caso D:\oracle\product\10.2.0\client_1\bin
```

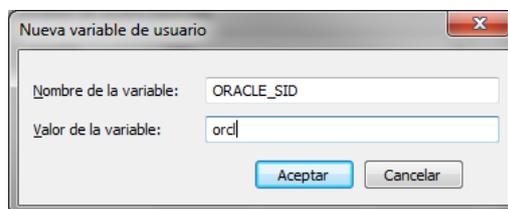


Gráfico N° 47: Variables de entorno del sistema

Fuente: Investigadora

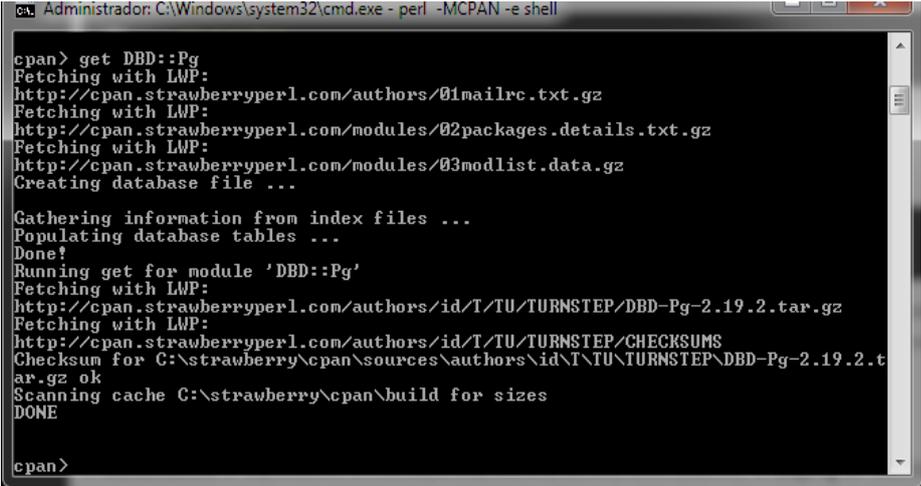
- Variables de entorno para conexión PostgreSQL

```
POSTGRES_HOME=C:\Program Files\PostgreSQL\9.0
POSTGRES_INCLUDE=C:\Program Files\PostgreSQL\9.0\include
POSTGRES_LIB=C:\Program Files\PostgreSQL\9.0\lib
```

2.Instalar distribución de Perl 5.6 o superior, instalar además los módulos DBI, DBD::Oracle, DBD::Pg.

Ejecutamos el comando de actualización de módulos de Perl desde la ventana de comandos de Windows para descargar e instalar el módulo Pg (para conectarse con PostgreSQL)

```
Perl -MCPAN -e Shell  
Get DBD::Pg;
```



```
ca. Administrador: C:\Windows\system32\cmd.exe - perl -MCPAN -e shell  
cpan> get DBD::Pg  
Fetching with LWP:  
http://cpan.strawberryperl.com/authors/01mailrc.txt.gz  
Fetching with LWP:  
http://cpan.strawberryperl.com/modules/02packages.details.txt.gz  
Fetching with LWP:  
http://cpan.strawberryperl.com/modules/03modlist.data.gz  
Creating database file ...  
Gathering information from index files ...  
Populating database tables ...  
Done!  
Running get for module 'DBD::Pg'  
Fetching with LWP:  
http://cpan.strawberryperl.com/authors/id/T/TU/TURNSTEP/DBD-Pg-2.19.2.tar.gz  
Fetching with LWP:  
http://cpan.strawberryperl.com/authors/id/T/TU/TURNSTEP/CHECKSUMS  
Checksum for C:\strawberry\cpan\sources\authors\id\T\TU\TURNSTEP\DBD-Pg-2.19.2.t  
ar.gz ok  
Scanning cache C:\strawberry\cpan\build for sizes  
DONE  
cpan>
```

Gráfico N° 48: Descarga de módulo DBD:Pg

Fuente: Investigadora

- Compilamos el código descargado con el comando *makefile* y luego el comando *dmake* dentro del directorio “CPAN/build” en nuestra ruta de Perl. La versión descargada fue DBD-Pg-2.192.

```

C:\strawberry\cpan\build\DBD-Pg-2.19.2-HTP1T0>dmake
cp testme.tmp.pl blib\lib\DBD\testme.tmp.pl
cp lib\Bundle\DBD\Pg.pm blib\lib\Bundle\DBD\Pg.pm
cp Pg.pm blib\lib\DBD\Pg.pm
C:\strawberry\perl\bin\perl.exe -p -e "s/^DRIVER~/Pg/g; s/^do</dontdo</ C:\st
rawberry\perl\vendor\lib\auto\DBI\Driver.xst > Pg.xsi
C:\strawberry\perl\bin\perl.exe C:\strawberry\perl\lib\ExtUtils\xsubpp -typemap
C:\strawberry\perl\lib\ExtUtils\typemap Pg.xs > Pg.xsc && C:\strawberry\perl\b
in\perl.exe -MExtUtils::Command -e "mv" -- Pg.xsc Pg.c
gcc -c -I"C:\Program Files\PostgreSQL\9.0\include" -IC:\strawberry\perl\vendor\
lib\auto\DBI -s -O2 -DWIN32 -DHAVE_DES_FCRYPT -DUSE_SITECUSTOMIZE -DPERL_IMP
LICIT_CONTEXT -DPERL_IMPLICIT_SYS -fno-strict-aliasing -DPERL_MSUCRT_READFIX -DP
GLIBVERSION=90007 -DPGDEFPORT=5432 -s -O2 -DUERSION="2.19.2" -DXS_UER
SION="2.19.2" "-IC:\strawberry\perl\lib\CORE" Pg.c
In file included from C:/strawberry/perl/lib/CORE/win32thread.h:4,
from C:/strawberry/perl/lib/CORE/perl.h:2637,
from C:/strawberry/perl/vendor/lib/auto/DBI/DBIXS.h:19,
from Pg.h:59,
from Pg.xs:13:
C:/strawberry/perl/lib/CORE/win32.h:238:1: warning: "isnan" redefined

```

Gráfico N° 49: Compilación módulo DBD::Pg

Fuente: Investigadora

- Instalamos el módulo compilado

```

C:\strawberry\cpan\build\DBD-Pg-2.19.2-HTP1T0>dmake install
Files found in blib\arch: installing files in blib\lib into architecture depende
nt library tree
Installing C:\strawberry\perl\site\lib\auto\DBD\Pg\Pg.bs
Installing C:\strawberry\perl\site\lib\auto\DBD\Pg\Pg.dll
Installing C:\strawberry\perl\site\lib\Bundle\DBD\Pg.pm
Installing C:\strawberry\perl\site\lib\DBD\Pg.pm
Installing C:\strawberry\perl\site\lib\DBD\testme.tmp.pl
Appending installation info to C:\strawberry\perl\lib\perllocal.pod
C:\strawberry\cpan\build\DBD-Pg-2.19.2-HTP1T0>_

```

Gráfico N° 50: instalación módulo DBD::Pg

Fuente: Investigadora

- Descargamos el módulo dbd::Oracle

```

C:\Users\User>perl -MCPAN -e shell
Copyright (c) 2009 Microsoft Corporation. Reservados todos los derechos.
C:\Users\User>perl -MCPAN -e shell
cpan shell -- CPAN exploration and modules installation (v1.9453)
Enter 'h' for help.

cpan> get DBD::Oracle
Database was generated on Sun, 20 May 2012 06:17:20 GMT
Running get for module 'DBD::Oracle'
Fetching with LWP:
http://cpan.strawberryperl.com/authors/id/P/PY/PYTHIAN/DBD-Oracle-1.44.tar.gz
Fetching with LWP:
http://cpan.strawberryperl.com/authors/id/P/PY/PYTHIAN/CHECKSUMS
Checksum for C:\strawberry\cpan\sources\authors\id\P\PY\PYTHIAN\DBD-Oracle-1.44.
tar.gz ok
Scanning cache C:\strawberry\cpan\build for sizes
.....DONE
cpan> _

```

Gráfico N° 51: Descarga del módulo DBD::Oracle

Fuente: Investigadora

- Compilamos el módulo DBD-Oracle. La versión descargada fue DBD-Oracle-1.44-X9dx2z

```

C:\Windows\system32\cmd.exe
20/05/2012 13:37 <DIR> .
20/05/2012 13:37 <DIR> ..
20/05/2012 13:37 <DIR> DBD-Oracle-1.44-X9dx2z
20/05/2012 01:29 <DIR> DBD-Pg-2.19.2-HTP1T0
                0 archivos          0 bytes
                4 dirs 57.097.912.320 bytes libres

C:\strawberry\cpan\build>cd DBD-Oracle-1.44-X9dx2z
C:\strawberry\cpan\build\DBD-Oracle-1.44-X9dx2z>perl makefile.pl

```

Gráfico N° 52: Compilación módulo DBD::Oracle

Fuente: Investigadora

```

C:\Windows\system32\cmd.exe
20/05/2012 13:37 <DIR> .
20/05/2012 13:37 <DIR> ..
20/05/2012 13:37 <DIR> DBD-Oracle-1.44-X9dx2z
20/05/2012 01:29 <DIR> DBD-Pg-2.19.2-HTP1T0
                0 archivos          0 bytes
                4 dirs 57.097.912.320 bytes libres

C:\strawberry\cpan\build>cd DBD-Oracle-1.44-X9dx2z
C:\strawberry\cpan\build\DBD-Oracle-1.44-X9dx2z>dmake

```

Gráfico N° 53: creación de ejecutable

Fuente: Investigadora

- Compilamos y ejecutamos el test del módulo DBD-Oracle.

```

C:\Windows\system32\cmd.exe - dmake test
C:\strawberry\cpan\build\DBD-Oracle-1.44-X9dx2z>dmake test
C:\strawberry\perl\bin\perl.exe "-MExtUtils::Command:MM" "-e" "test_h
'blib\lib', 'blib\arch'" t/*.t
t/000-report-versions.t ... # Testing with Perl 5.008009, C:\STRAWB~1\
perl.exe
# Carp version is 1.10
t/000-report-versions.t ... 1/? # Config version is undefined
# DBI version is 1.609
# Data::Dumper version is 2.125
# Devel::Peek version is 1.04
# DynaLoader version is 1.09
# Encode version is 2.39
# Exporter version is 5.63
# ExtUtils::MakeMaker version is 6.56
# Math::BigInt version is 1.89

```

- Verificamos la correcta ejecución del test del módulo DBD-Oracle.

```

t/50cursor.t ..... ok
t/51scroll.t ..... ok
t/55nested.t ..... ok
t/56embbeded.t ..... ok
t/58object.t ..... ok
t/60reauth.t ..... skipped: ORACLE_USERID_2 not defined.
t/70meta.t ..... ok
t/80ora_charset.t ..... ok
t/rt13865.t ..... ok
t/rt74753-utf8-encoded.t .. skipped: database character set is not Uni
All tests successful.
Files=36, Tests=2601, 890 wallclock secs < 1.29 usr + 0.30 sys = 1.5
Result: PASS
C:\strawberry\perl\bin\perl.exe "-Iblib\lib" "-Iblib\arch" test.pl
C:\strawberry\cpan\build\DBD-Oracle-1.44-x9dx2z>

```

Gráfico N° 54: Ejecución de Test de módulo Oracle

Fuente: Investigadora

- Se instala el módulo dbd::Oracle

```

ca. Administrador: C:\Windows\system32\cmd.exe
C:\strawberry\perl\bin\perl.exe "-Iblib\lib" "-Iblib\arch" test.pl
C:\strawberry\cpan\build\DBD-Oracle-1.44-x9dx2z>dmake install
Files found in blib\arch: installing files in blib\lib into architect
nt library tree
Installing C:\strawberry\perl\site\lib\auto\DBD\Oracle\mk.pm
Installing C:\strawberry\perl\site\lib\auto\DBD\Oracle\Oracle.dll
Appending installation info to C:\strawberry\perl\lib\perllocal.pod
C:\strawberry\cpan\build\DBD-Oracle-1.44-x9dx2z>

```

Gráfico N° 55: Instalación de módulo Oracle

Fuente: Investigadora

3. Instalamos la herramienta de migración ora2Pg

Una vez instalados todos los requerimientos se instala la herramienta Ora2Pg, se copia el archivo fuente descargado de la web, en la carpeta donde se descargaron los módulos de perl en “CPAN/Build”.

- Compilamos por medio del comando *makefile*

```
C:\Windows\system32\cmd.exe
C:\strawberry\cpan\build\ora2pg-8.9>cd..
C:\strawberry\cpan\build>cd ora2pg-8.9
C:\strawberry\cpan\build\ora2pg-8.9>perl makefile.pl
"which" no se reconoce como un comando interno o externo,
programa o archivo por lotes ejecutable.
"chmod" no se reconoce como un comando interno o externo,
programa o archivo por lotes ejecutable.
Can't find string terminator '"' anywhere before EOF at -e line 1.
Checking if your kit is complete...
Looks good
Writing Makefile for Ora2Pg::PLSQL
Writing Makefile for Ora2Pg
Done...

Now type 'make && make install'

C:\strawberry\cpan\build\ora2pg-8.9>dmake
cp Ora2Pg.pm blib\lib\Ora2Pg.pm
cp PLSQL.pm ..\blib\lib\Ora2Pg\PLSQL.pm
C:\strawberry\cpan\build\ora2pg-8.9>dmake install
```

Gráfico N° 56: Compilación de Ora2Pg

Fuente: Investigadora

- Finalmente instalamos Ora2Pg

```
C:\Windows\system32\cmd.exe
"chmod" no se reconoce como un comando interno o externo,
programa o archivo por lotes ejecutable.
Can't find string terminator '"' anywhere before EOF at -e line 1.
Checking if your kit is complete...
Looks good
Writing Makefile for Ora2Pg::PLSQL
Writing Makefile for Ora2Pg
Done...

Now type 'make && make install'

C:\strawberry\cpan\build\ora2pg-8.9>dmake
cp Ora2Pg.pm blib\lib\Ora2Pg.pm
cp PLSQL.pm ..\blib\lib\Ora2Pg\PLSQL.pm
C:\strawberry\cpan\build\ora2pg-8.9>dmake install
Installing C:\strawberry\perl\site\lib\Ora2Pg.pm
Installing C:\strawberry\perl\site\lib\Ora2Pg\PLSQL.pm
sh install_all.sh
"sh" no se reconoce como un comando interno o externo,
programa o archivo por lotes ejecutable.
dmake: Error code 129, while making 'install_all'
C:\strawberry\cpan\build\ora2pg-8.9>
```

Gráfico N° 57: Instalación de Ora2Pg

Fuente: Investigadora

4. Probar conexiones entre hosts que intervendrán en el proceso de migración

Al host servidor de Oracle *msping 192.168.20.100*

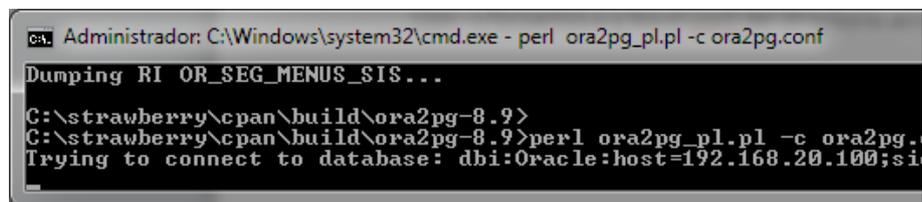
Probar la conexión al host postgresSQL *ping 192.168.20.50*

5. En el host donde se encuentra instalada la herramienta de migración y modificamos el archivo de configuración para especificar el tipo de información a obtener.

```
# Type of export. Values can be the
following keyword:
#     TABLE      Export tables
#     PACKAGE     Export packages
#     TRIGGERS
#     DATA       Export data from table
as INSERT statement
```

6. Obtener la extracción de la estructura de base de datos. Ejecutar la herramienta ora2Pg mediante el siguiente comando:

Perl ora2pg_pl.pl -c ora2pg.conf



```
C:\Administrador: C:\Windows\system32\cmd.exe - perl ora2pg_pl.pl -c ora2pg.conf
Dumping RI OR_SEG_MENUS_SIS...
C:\strawberry\cpan\build\ora2pg-8.9>
C:\strawberry\cpan\build\ora2pg-8.9>perl ora2pg_pl.pl -c ora2pg.conf
Trying to connect to database: dbi:Oracle:host=192.168.20.100;si
```

Gráfico N° 58: Instalación de Ora2Pg

Fuente: Investigadora

7. **En el directorio de exportación revisar los scripts obtenidos y abrirlos para su modificación en caso de ser necesario.**

Cambiar tipo de dato CBLOB a text

Cambiar tipo de dato varchar2 a text

8. Abrir el cliente postgresQL y ejecutar los scripts de creación de la estructura

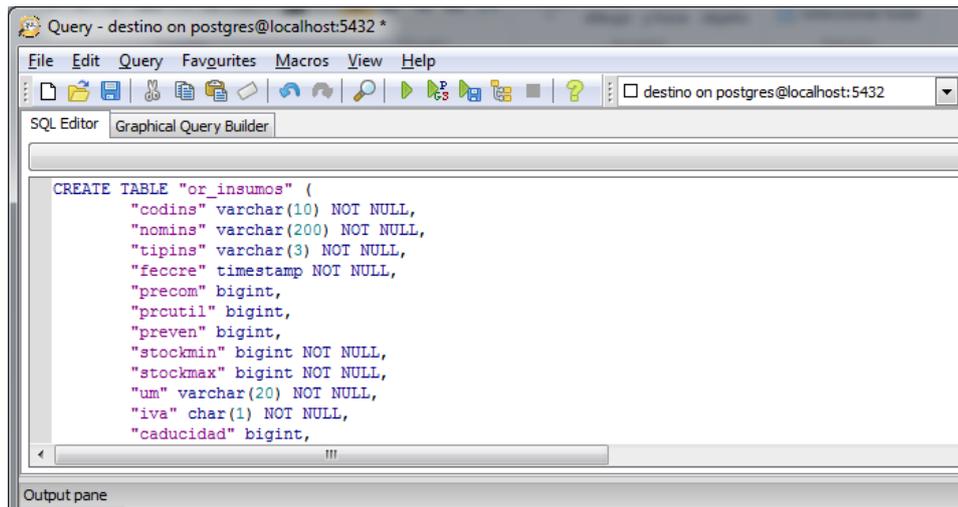


Gráfico N° 59: Ejecución de scripts en PostgreSQL

Fuente: Investigadora

9. Realizar un control de calidad y comparativa de la estructura priorizando las tablas de información crítica.

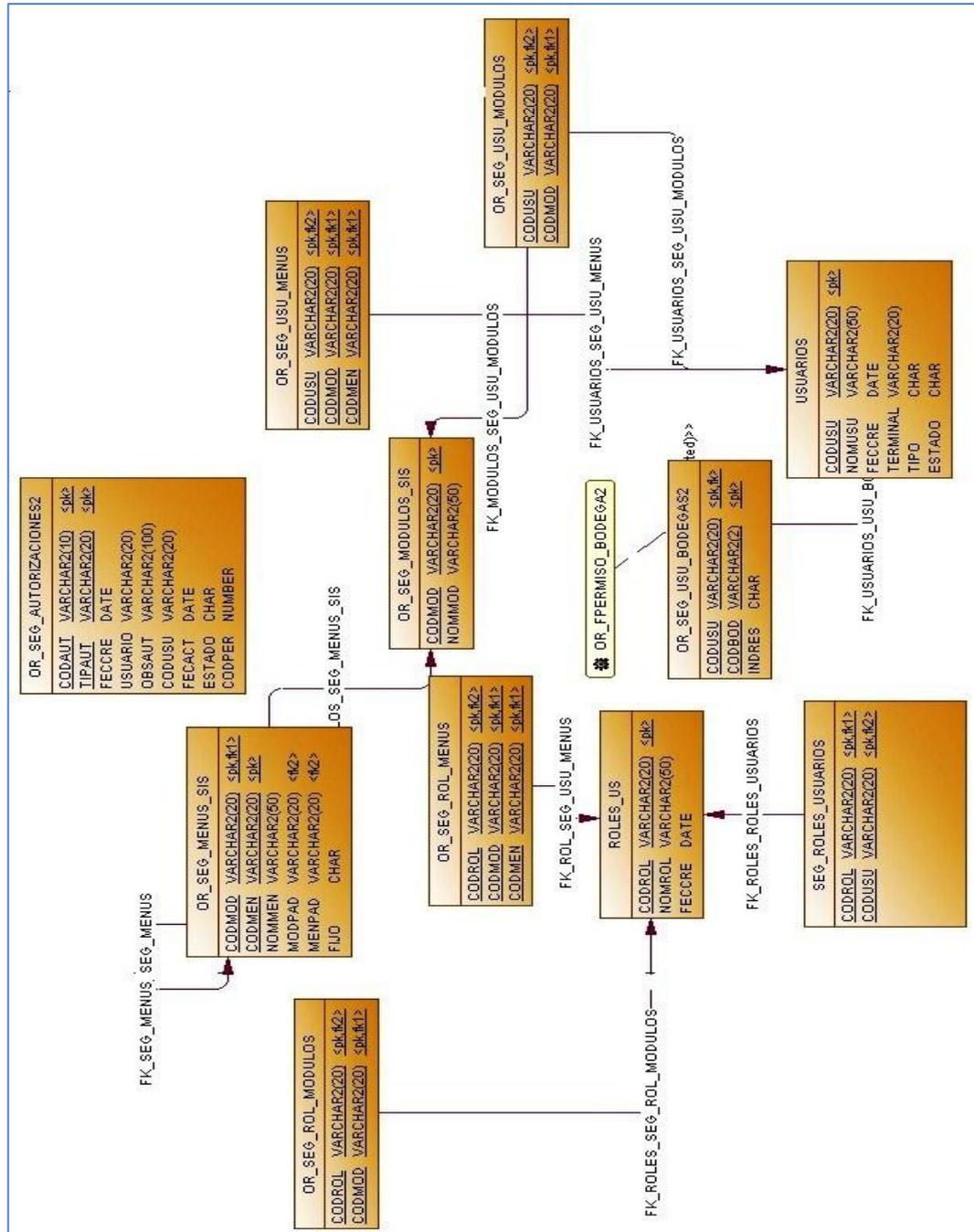
10 Extraer los datos de las tablas, previamente modificamos el archivo .conf de ora2pg para señalar que se va a extraer los datos.

```
# Type of export. Values can be the following  
keyword:  
DATA  
#Export data from table as INSERT statement
```

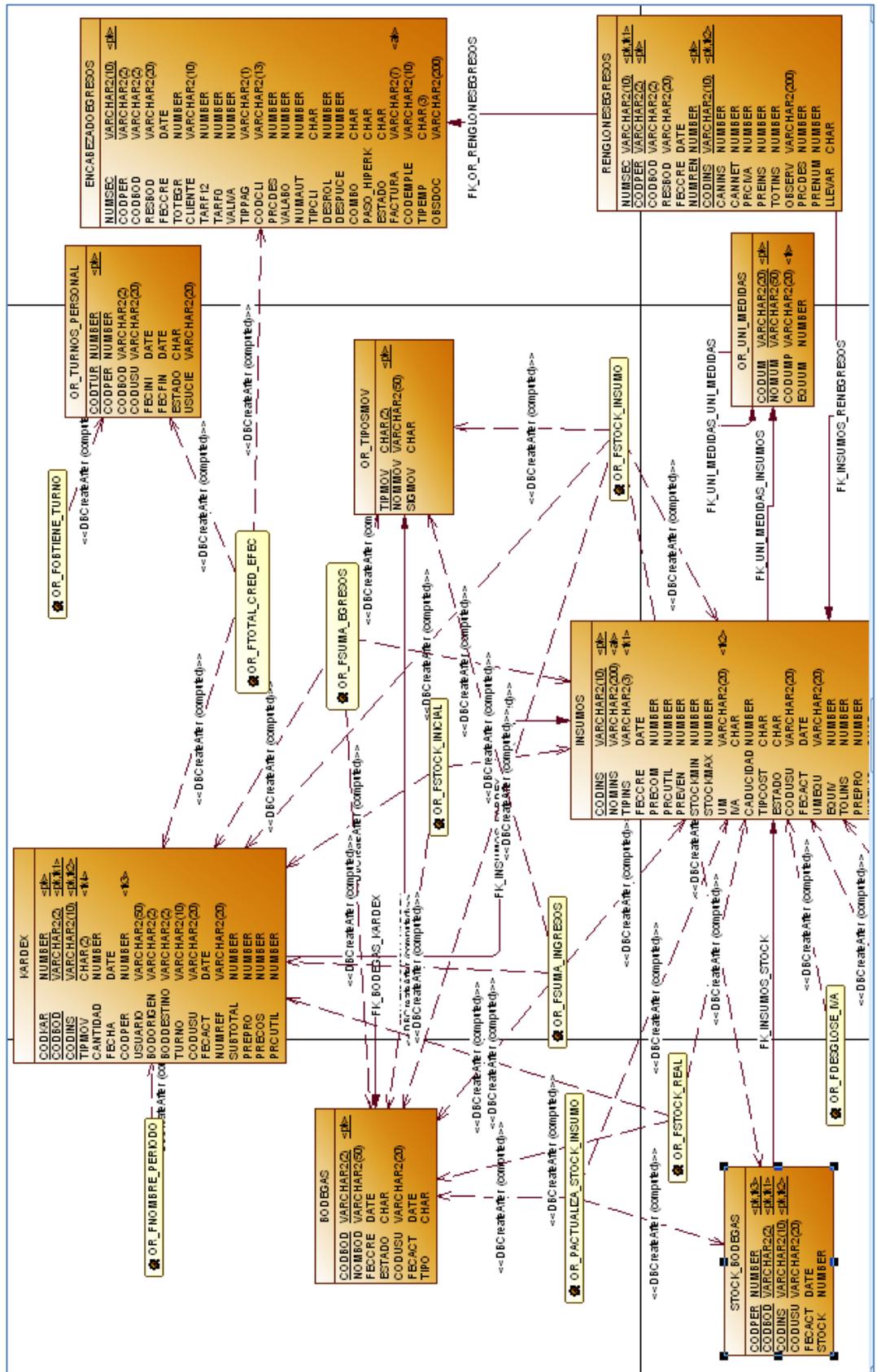
11. Abrir los scripts y modificarlos en caso de ser necesario para su inserción

ANEXO A6: Modelo de datos

- Módulo de usuarios



- Módulo de Inventarios(1)



ANEXO A7: GUÍA DE LA ENTREVISTA PARCIALMENTE ESTRUCTURADA APLICADA A GERENTE DE LA EMPRESA VSYSTEMS

<p>Nº.</p> <p>NOMBRE DE LA INSTITUCIÓN.....</p> <p>ENTREVISTADO.....</p> <p>ENTREVISTADORA.....</p> <p>LUGARFECHA</p> <p>OBJETO DE ESTUDIO.....</p>	
PREGUNTAS	INTERPRETACIÓN-VALORACIÓN
<p>1. ¿Cuáles los beneficios que Ud. desea obtener al aplicar este proceso de migración en su empresa?</p>	
<p>2. ¿Existe en su empresa un análisis de información que establezca y determine cuál es la información crítica?</p>	
<p>3. Cuáles son las causas que dan origen a este proceso de migración de datos en su empresa?</p>	
<p>4. Describa como ha sido su crecimiento empresarial y como esto ha repercutido en la información que administra la empresa?</p>	

ANEXO A8:

GUÍA DE LA ENTREVISTA PARCIALMENTE ESTRUCTURADA APLICADA AL ADMINISTRADOR DE BASE DE DATOS

<p>Nº.</p> <p>NOMBRE DE LA INSTITUCIÓN.....</p> <p>ENTREVISTADO.....</p> <p>ENTREVISTADORA.....</p> <p>LUGARFECHA</p> <p>OBJETO DE ESTUDIO.....</p>	
PREGUNTAS	INTERPRETACIÓN-VALORACIÓN
<p>Describa brevemente los principales procedimientos y operaciones de los sistemas de información en relación con la base de datos.</p>	
<p>Describa el estándar o nomenclatura que es empleada en la estructura de la base de datos.</p>	
<p>Describa las técnicas que utilizan para validar la información de su base de datos actual</p>	
<p>Explique las características más importantes de su gestor actual de base de datos.</p>	

ANEXO A8: INSTRUMENTOS PARA LA ENCUESTA

ENCUESTA DIRIGIDA A ADMINISTRADORES, OPERADORES Y USUARIOS DE BASE DE DATOS DE LA EMPRESA VSYSTEMS

OBJETIVO: Determinar el estado inicial de la Base de datos y conocimiento de procesos de migración de datos.

Estamos trabajando en un estudio que servirá para elaborar una tesis profesional acerca de la creación de una metodología de migración de datos que asegure conservación de integridad y consistencia.

Sus respuestas serán confidenciales y anónimas.

DATOS GENERALES:

Fecha de la Encuesta.....

DATOS ESPECÍFICOS: Marque con X en el paréntesis de su elección

N.	PREGUNTAS	RESPUESTAS	COD.
1	El proceso de migración que se va a implementar en la empresa es causado por:	<ul style="list-style-type: none"> • Innovación Tecnológica • Ahorro de costos de licencias • El gestor actual no cumple con requerimientos de información. 	1. () 2. () 3. ()
2	¿Considera que la estructura de la base de datos es?	<ul style="list-style-type: none"> • Muy compleja • Compleja. • Simple 	1. () 2. () 3. ()
3	Conoce o disponen de una metodología para aplicar la migración	<ul style="list-style-type: none"> • SI • NO 	1. () 2. ()
4	¿El tamaño de la base de datos es?	<ul style="list-style-type: none"> • Grande >1000 tablas • Mediana (100-999) • Pequeña menos de 100 tablas 	1. () 2. () 3. ()

5	Cuál de los siguientes aspectos considera usted que es mayor gravedad al realizar un proceso de migración.	<ul style="list-style-type: none"> • Pérdida de información. • Pérdida de integridad de datos. • Incompatibilidad entre tipo de datos. 	<p>1. ()</p> <p>2. ()</p> <p>3. ()</p>
6	La base de datos tiene documentación sobre su diseño y estructura	<ul style="list-style-type: none"> • SI. • NO 	<p>1. ()</p> <p>2. ()</p>
7	Realiza análisis de información y conoce cuál es la información crítica de su base de datos?	<ul style="list-style-type: none"> • SI • NO 	<p>1.()</p> <p>2.()</p>
8	Se realiza algún procedimiento periódico para depuración y ejecutar control de calidad sobre los datos.	<ul style="list-style-type: none"> • SI. • NO 	<p>1.()</p> <p>2.()</p>
9	Las aplicaciones que interactúan con la base de datos consiste en	<ul style="list-style-type: none"> • Software desarrollado por la empresa • Software comprado sin personalización • Software comprado con personalización 	<p>1. (x)</p> <p>2. ()</p> <p>3. ()</p>
10	El crecimiento de la base de datos en el último año se encuentra dentro de los siguientes porcentajes	<ul style="list-style-type: none"> • Menos del 25% • Entre 25% y 50% • Entre 50% y 75% • Más del 75% 	<p>1. ()</p> <p>2. ()</p> <p>3. ()</p> <p>4. ()</p>
11	Luego de haber desarrollado algún proceso migratorio su base de datos presentó inconsistencias graves.	<ul style="list-style-type: none"> • SI. • NO 	<p>1. ()</p> <p>2. ()</p>

Gracias por su colaboración.