

Universidad de las Ciencias Informáticas

Facultad 3



Título: Mercados de Datos para los departamentos Verificación Fiscal, Quejas Peticiones y Denuncias, Dictámenes y Revisiones de Causas Penales de la Fiscalía General de la República.

**TRABAJO DE DIPLOMA PARA OPTAR POR EL TÍTULO DE INGENIERO EN CIENCIAS
INFORMÁTICAS**

Autores: Yordanka Parellada Batista

Tutor: Ing. Manuel Álvarez Alonso

Co-Tutores: Ing. José Carlos Pupo Acosta

Ing. Yeleny Almora Galvez

La Habana, junio de 2014



"Estar preparado es importante, saber estarlo es aún más, pero aprovechar el momento adecuado es la clave de la vida."

Arthur Schnitzler

Declaración de autoría

Yo Yordanka Parellada Batista declaro ser la autora del trabajo de diploma con título: Mercados de Datos para los departamentos Verificación Fiscal, Quejas, Peticiones y Denuncias, Dictámenes y Revisiones de Causas Penales de la Fiscalía General de la República y admito los derechos patrimoniales del mismo con carácter exclusivo a la Universidad de las Ciencias Informáticas, específicamente a la Facultad 3.

Para que así conste, firmo el presente a los ____ días del mes de _____ del año 2014.

Firma de la autora
Yordanka Parellada Batista

Firma del tutor
Manuel Álvarez Alonso

Firma del tutor
José Carlos Pupo Acosta

Firma de la tutora
Yeleny Almora Galvez

Datos de contacto

TUTORES:

Ing. Manuel Álvarez Alonso

Universidad de las Ciencias Informáticas,
Ciudad de la Habana, Cuba

E-mail: malvareza@uci.cu

Ing. José Carlos Pupo Acosta

Universidad de las Ciencias Informáticas,
Ciudad de la Habana, Cuba

E-mail: jcpupo@uci.cu

Ing. Yeleny Almora Galvez

Universidad de las Ciencias Informáticas,
Ciudad de la Habana, Cuba

E-mail: yalmora@uci.cu

Dedicatoria

A toda mi familia, que siempre me han apoyado incondicionalmente recordándome que puedo contar con ellos en todo momento, a pesar de ser una familia pequeña y con muchos defectos siempre estamos ahí los unos para los otros, pero especialmente a mi papá Ibrahín, que ha apoyado cada decisión que he tomado en mi vida aun cuando no estuviera de acuerdo, a mi mamá Maricel, quien entre muchas otras cosas me enseñó a perdonar, porque cuando no lo hacemos sin saberlo somos los principales perjudicados, a mis hermanos Ibrahín y José Carlos que son muy importantes para mí, a mis abuelas Belkis y Laude que no han hecho más que velar por mí y consentirme mucho y a mi novio Daniel Arturo y su familia que me han apoyado mucho en los últimos tres años.

A la Revolución Cubana, sin la cual no hubiera sido posible este acontecimiento, en especial a su máximo líder Fidel Castro Ruz.

Agradecimientos

Primeramente quiero agradecerle a mi familia, pero muy especialmente a mi papá y a mi mamá, todo esto se lo debo a ellos y es gracias a ellos que hoy puedo decir que lo logré, porque siempre me dieron la fuerza que necesitaba, me dijeron las palabras justas en el momento justo, siempre mostrándome la verdad y no lo que yo quería ver, por eso y por todo mil gracias. También quiero agradecer a mis hermanos Ibrahín y José Carlos, que más que hermanos son mis amigos y representan una parte fundamental en mi vida. A mis abuelos que se han desvelado, me han cuidado, consentido, querido y apoyado desde que di mis primeros pasos, me enseñaron a ser fuerte y a seguir adelante con la frente en alto a pesar de cualquier situación que se pudiera presentar en la vida. A mi novio Daniel por haber estado conmigo durante los últimos tres años y apoyarme en todo al igual que su familia, a todos les agradezco con el cariño y dedicación que me adoptaron y por hacerme sentir un miembro más de su hermosa familia. A mis primos, que aunque no somos muchos nos queremos y somos muy buenos amigos. A mis tías y tíos a los cuales quiero mucho y sé que ellos a mí también. A mi mejor amiga Anabel, que más que una amiga es como mi hermana, hemos vivido muchas cosas juntas y espero que siga siendo así. A los buenos amigos que hice en la universidad, en especial a Fidel, Alberto, Yamilka, Dania, Pedro, Leo, Aymé, Odeysis, Anidey, Yandry, Jeandy, Carlos Miguel, pero especialmente al F4, que han estado conmigo en las buenas y en las malas, nos hemos reído juntos, nos hemos apoyado unos a otros y hemos llorado unos con otros demostrándonos cada día que el valor de una bonita amistad es incalculable. A todos mis compañeros de aula, les agradezco infinitamente por preocuparse por mí, por correr conmigo siempre que fue necesario y por cuidarme, a Patricio, Orlando, Ariam, Yaimaris, Sary, Kenia, Yasel, Yoelvis, Leonardo Ricardo, Rayner, Víctor, Yoalvis, Javier A, Javier B y Javier S, Ledian, Osmail, Lanet, Nilson, Yesica, Wanda a todos muchas gracias, realmente me siento en deuda con todos ustedes porque desde el primer momento que entré a este grupo me acogieron y cuidaron en cada momento, además por hacer que estos años fueran más divertidos, a los profesores que me han dado clases y han estado pendientes siempre de mi salud y me han apoyado, a mis tutores, por haberme guiado durante este curso. A la revolución cubana por darme la oportunidad de forjarme como ingeniera. A todos los que de una forma u otra han estado ahí para mí mil gracias, siempre lo tendré presente.

Resumen

La información se ha convertido en estos días en un elemento, más que necesario, imprescindible para la toma de decisiones en las empresas. Solo aquellas empresas que son capaces de tomar decisiones correctas a tiempo pueden adaptarse al agitado mundo en que vivimos, pero para ello necesitan información con calidad y rapidez. A esta realidad no escapa el mundo jurídico y en nuestro país esta situación se vigencia en las distintas instancias de la Fiscalía General de la República.

En la presente investigación se expone la creación de tres Mercados de Datos para los departamentos Verificación Fiscal, Quejas, Peticiones y Denuncias, Dictámenes y Revisiones de causas Penales. En aras de desarrollarlos se realizaron entrevistas a los analistas del proyecto Sistema de Informatización de la Gestión de la Fiscalía para determinar las necesidades de información y profundizar en el proceso de negocio de la Fiscalía, además se realizó un estudio de los fundamentos teóricos. Un análisis detallado de las Bases de Datos operacionales, conjuntamente con los requisitos específicos del usuario, aplicando la metodología seleccionada, culminaron con el diseño de los Mercados de Datos. Se utilizaron en la construcción de las Bases de Datos de los Mercados, los procesos ETL (Extracción, Transformación y Carga) con la herramienta Pentaho Data Integration y en la carga de los cubos OLAP herramientas como Schema Workbench y Biserver-ce. Con esto se mejora y acelera el proceso de toma de decisiones en la Fiscalía General de la República, permitiendo elevar la efectividad de la organización.

Palabras claves: Almacén de Datos, Bases de Datos, Mercado de Datos, Toma de Decisiones.

ÍNDICE

Introducción	1
1 Capítulo 1. Fundamentos teóricos	6
1.1 Introducción	6
1.2 Almacenes de Datos.....	¡Error! Marcador no definido.
1.2.1 Características de los Almacenes de Datos.....	6
Ventajas y desventajas de los Almacenes de Datos	7
1.3 Mercado de Datos	7
1.3.1 Características de los Mercados de Datos.....	8
1.4 Arquitectura de un Almacén de Datos.....	8
1.5 Metodología de Desarrollo	9
1.6 Modelos de datos	10
1.7 Base de Datos multidimensional.....	11
1.7.1 Variantes de modelado	12
1.8 Extracción, Transformación y Carga de los datos	14
1.9 Herramienta de modelado	15
1.9.1 Visual Paradigm 8.0.....	15
1.10 Sistema Gestor de Base de Datos.....	15
1.10.1 PostgreSQL 9.1	16
1.10.2 Administrador de Base de Datos. PgAdmin III 1.14.0	16
1.11 Herramientas para la Extracción, Transformación y Carga de datos.....	17
1.11.1 Pentaho Data Integration 4.2.1	17
1.11.2 Mondrian Schema WorkBench	18

1.11.3	Pentaho BI server 4.8.0	18
1.12	Herramientas para las pruebas.....	19
1.12.1	Apache JMeter 2.9.....	19
1.12.2	Datanamic Data Generator for PostgreSQL 5.0.....	19
1.13	Conclusiones del capítulo	20
2	Capítulo 2. Diseño e implementación de los Mercados de Datos.....	21
2.1	Introducción.....	21
2.2	Análisis de las necesidades de información.....	21
2.3	Análisis del estado de las fuentes de datos	24
2.4	Metodología de Kimball	24
2.4.1	Planificación.....	24
2.4.2	Análisis de requerimientos	25
2.4.3	Matriz bus o matriz dimensional	27
2.4.4	Modelado Dimensional.....	30
2.4.5	Elegir el proceso de negocio	30
2.4.6	Establecer el nivel de granularidad.....	30
2.4.7	Identificación de las dimensiones	38
2.4.8	Identificación de las tablas de hechos	41
2.4.9	Identificación de atributos de dimensiones y tablas de hechos.....	42
2.4.10	Patrones de diseño de Bases de Datos para los MD.....	46
2.4.10.1	Patrón de llaves subrogadas.....	46
2.4.11	Estandarización del código.....	47
2.4.12	Tipo de modelo lógico del Mercado de Datos	48
2.4.13	Arquitectura de los Mercados de Datos.....	48

2.4.14	Modelo lógico de los datos	48
2.4.15	Modelo físico de los datos	49
2.4.16	Prueba del modelo	52
2.4.17	Diseño de los sistemas de extracción, transformación y carga de los datos.....	53
2.5	Implementación de los trabajos	55
2.6	Diseño de las dimensiones y cubo de información.....	55
2.7	Visualización de los resultados	56
2.8	Conclusiones del capítulo	56
3	Capítulo 3. Pruebas	58
3.1	Introducción	58
3.2	Pruebas	58
3.3	Pruebas de carga y estrés	58
3.4	Pruebas de Volumen y Carga	60
3.5	Conclusiones del capítulo	63
	Conclusiones Generales.....	64
	Recomendaciones	65
	Referencias bibliográficas	66
	Anexos.....	68

Índice de Figuras

Figura 1. Enfoque descendente o top-down definido por Bill Inmon.....	9
Figura 2. Enfoque ascendente o bottom-up definido por Ralph Kimbal.....	10
Figura 3. Esquema en estrella	13
Figura 4. Esquema en copo de nieve.....	13
Figura 5. Esquema de constelación de hechos.....	14
Figura 6. Arquitectura del sistema.....	48
Figura 7. Modelo lógico de datos para el subsistema PDC.	49
Figura 8. Modelo físico de datos para el subsistema CLEP.	50
Figura 9. . Modelo físico de datos para el subsistema VF.	51
Figura 10. Modelo físico de datos para el subsistema PDC.	52
Figura 11. Fragmento del modelo conceptual.	53
Figura 12. Transformación para la dimensión “dim_tiempo”.....	54
Figura 13. Transformación para el hecho “hecho_rollo_queja”.....	54
Figura 14. Transformación para el hecho “hecho_rollo_fgr”	54
Figura 15. Transformación para el hecho “hecho_acciones.	54
Figura 16. Representación en el navegador del cubo Rollo_FGR.....	56
Figura 17. Resultado de la prueba de carga simulando cantidad de usuarios de la Fiscalía General.....	59
Figura 18. Resultado de la prueba de carga simulando cantidad de usuarios de la Fiscalía Provincial.	59
Figura 19. Resultado de la prueba de Carga en Datanamic Data Generator para el mercado del subsistema PDC.....	61
Figura 20. Resultado de la prueba de carga simulando cantidad de usuarios de la Fiscalía General.....	62
Figura 21. Resultado de la prueba de carga simulando cantidad de usuarios de la Fiscalía Provincial.	62

Índice de tablas

Tabla 1. Temas analíticos. Subsistema CLEP.	25
Tabla 2. Temas analíticos. Subsistema VF.....	26
Tabla 3. Temas analíticos. Subsistema PDC.....	27
Tabla 4. Nivel de granularidad para el subsistema CLEP.	34
Tabla 5. Nivel de granularidad para el subsistema PDC.	36
Tabla 6. Nivel de granularidad para el subsistema VF.....	38
Tabla 7. Atributos de las tablas dimensiones del subsistema CLEP.	43
Tabla 8. Atributos de las tablas dimensiones del subsistema PDC.....	44
Tabla 9. Atributos de las tablas dimensiones del subsistema VF.....	44
Tabla 10. Atributos de las tablas hechos del subsistema CLEP.....	45
Tabla 11. Atributos de las tablas hechos del subsistema VF.	45
Tabla 12. Atributos de las tablas hechos del subsistema PDC.	46

Introducción

En la sociedad actual, el mundo cursa por la era de la información. El poseer conocimiento se ha tornado un arma considerablemente poderosa, es obvio que muchos enfoquen sus esfuerzos en obtenerlo. (1) Muchas empresas, hoy en día, están inundadas de datos, pero lo que realmente necesitan es información y conocimiento; pues la competencia depende en gran medida del análisis de los datos del negocio, el conocimiento del mercado y la toma de decisiones oportunas.

Todas las empresas necesitan que la información sea eficaz y confiable en el proceso de toma de decisiones, con el fin de extraer conocimiento. Es un hecho que la competencia crece constantemente, así que las decisiones que se tomen deben ser aceleradas y de calidad, es por esto que se requieren herramientas que contribuyan con la minimización del tiempo, para analizar gran cantidad de datos con una mayor velocidad y precisión. Basado en esta información comprensible, detallada y relevante se deriva el conocimiento, necesario para lograr y sostener ventaja competitiva. El poseer conocimientos adecuados significa tener respuestas acertadas y realizar decisiones estratégicas para la ejecución de la empresa. Pero las tareas de recolectar, procesar, limpiar y transformar los datos en información necesaria para la toma de decisiones no es una tarea sencilla.

Con la informatización de la sociedad y dentro de éstas las empresas, ha crecido a nivel mundial la capacidad de generación y almacenamiento de la información, que no puede ser analizada por los métodos tradicionales existentes, mientras mayor es la capacidad para almacenar más y más datos, mayor es la incapacidad para extraer información realmente útil de éstos en las empresas. Mucha información importante, queda sepultada y disgregada, y los sistemas existentes no están preparados para el nuevo reto.

La era de la información y sus ventajas, ha sido un motivo del que hablar para muchos, al existir nuevas posibilidades se acortan las distancias y crecen los beneficios para quienes tienen acceso a las grandes cantidades de datos. Sin embargo, manejar grandes volúmenes de información de forma óptima se ha convertido en un reto importante de principios de siglo. La competencia en el nuevo ambiente corporativo donde todo está marcado por la información, el conocimiento del mercado y la toma de decisiones, es muy importante saber dónde y cómo se organiza toda la información.

En la actualidad, el centro de atención de las grandes instituciones, para lograr una ventaja competitiva, se ha enfocado en el análisis de los datos organizacionales, a través de diversas herramientas que interactúan con un Almacén de Datos. Los Almacenes de Datos surgen por la necesidad de resolver problemas de análisis de grandes masas de información. Ellos aseguran una vista única de los datos, que pueden provenir de diversas fuentes. Gracias a esto, los usuarios finales no se ven en la necesidad de aprender y utilizar múltiples sistemas de acceso y manipulación de los datos (2)

La Universidad de las Ciencias Informáticas (UCI) tiene la misión de informatizar la sociedad cubana y como resultado, en mutuo acuerdo con la Fiscalía General de la República (FGR), se ha creado el proyecto Fiscalía, que desarrolla el Sistema de Gestión Fiscal (SIGEF), encargado de informatizar los procesos de la Fiscalía.

Entre los departamentos del SIGEF se encuentran Verificación Fiscal, Quejas Peticiones y Denuncias, Dictámenes y Revisiones de Causas Penales que tienen como objetivo principal informatizar los procesos del mismo nombre en los distintos niveles de las FGR. La dirección de la FGR y la de estos departamentos necesitan poseer un repositorio de datos histórico que permita un acceso rápido al mismo. El sistema transaccional existente no facilita la búsqueda de información histórica con agilidad, ni la realización de comparaciones con la totalidad de los datos significativos en un determinado intervalo en el tiempo, debido a la forma en que está organizada dicha información; esta situación atenta contra la efectividad del proceso de toma de decisiones de la dirección de los departamentos Verificación Fiscal, Quejas Peticiones y Denuncias, Dictámenes y Revisiones de Causas Penales y los de la FGR.

De la situación problemática antes expuesta se extrae como **Problema científico**: ¿Cómo contribuir a la toma de decisiones en los departamentos Verificación Fiscal, Quejas Peticiones y Denuncias, Dictámenes y Revisiones de Causas Penales de la Fiscalía General de la República?

El **objeto de estudio** está orientado a los Sistemas de Apoyo a la toma de Decisiones.

El **campo de acción** es la Gestión de los Mercados de Datos.

Para solucionar el problema de la investigación se identifica como **objetivo general**: Desarrollar tres Mercados de Datos para los departamentos Verificación Fiscal, Quejas Peticiones y Denuncias, Dictámenes y Revisiones de Causas Penales de la Fiscalía General de la República que contribuya a la toma de decisiones de los directivos de estas áreas.

Para cumplir el objetivo general del trabajo se trazan los siguientes **objetivos específicos**:

- Realizar el marco teórico referencial de la investigación.
- Identificar las necesidades de información.
- Diseñar los Mercados de Datos.
- Realizar los procesos de extracción, transformación y carga.
- Realizar pruebas a la solución.

Para dar cumplimiento a los objetivos específicos planteados se proponen las siguientes **tareas de la investigación**:

- Estudio del estado del arte sobre los principales sistemas de soporte a la toma de decisiones y sistemas de software que realizan procesos similares.
 - Estudiar los temas relacionados a las mejores prácticas en el desarrollo de Almacén de Datos.
 - Definir la metodología a utilizar en el desarrollo.
- Obtención del análisis de los principales requerimientos de información en los departamentos Verificación Fiscal, Quejas Peticiones y Denuncias, Dictámenes y Revisiones de Causas Penales.
- Definición de la granularidad del proceso del negocio.
- Definición de las dimensiones de los Mercados de Datos.
- Definición de los hechos mensurables asociados a las dimensiones definidas.
- Estructuración del modelo dimensional.
- Transformación del modelo lógico dimensional al diseño físico.
- Definición de la estrategia de Extracción, Transformación y Carga de los datos.

- Implementación de los Mercados de Datos.
- Validación la solución implementada mediante pruebas de software.

Lo que se espera obtener con este trabajo son tres Mercados de Datos para el análisis de los datos generados en los departamentos Verificación Fiscal, Quejas Peticiones y Denuncias, Dictámenes y Revisiones de Causas Penales del Sistema de Gestión de la Fiscalía. Para esto se determinaron algunos métodos científicos que nos van a servir como apoyo para la realización de la investigación.

Los **métodos científicos** son los procedimientos usados para estudiar la realidad existente con el propósito de descubrir su esencia y sus relaciones, los que se utilizarán para darle solución a los objetivos propuestos son:

Métodos teóricos:

Analítico – sintético: este método se utiliza en el estudio de las diferentes bibliografías para sintetizar el conocimiento aprendido.

Modelación: El modelo es una reproducción simplificada de la realidad que permite descubrir nuevas relaciones y cualidades del objeto de estudio. Este método se evidenciará en el diseño de los Mercados de Datos.

Métodos empíricos:

Simulación: Posibilitó probar las Bases de Datos de los Mercados desarrollados con datos artificiales para de esta manera verificar el correcto funcionamiento de los Mercados de Datos.

La presente investigación se encuentra desglosada en tres capítulos estructurados de la siguiente manera:

Capítulo 1: Fundamentación teórica sobre el desarrollo de los Mercados de Datos. En este capítulo se realiza una breve explicación de la metodología, técnicas y herramientas que se tuvieron en cuenta para dar solución al problema; así como los principales conceptos y características de las mismas.

Capítulo 2: Diseño e implementación de los Mercados de Datos y sus etapas. En este capítulo se muestran aspectos relacionados con el modelado y diseño de la solución propuesta y se realiza la implementación de

las reglas de transformación de la Base de Datos relacional a los Mercados de Datos, así como las funciones para mostrar la información.

Capítulo 3: Prueba. En este capítulo se realizarán un conjunto de pruebas con el fin de garantizar el funcionamiento de los Mercados de Datos, dichas pruebas son las pruebas de Volumen y Carga y las pruebas de Carga y Estrés.

1 Capítulo 1. Fundamentos teóricos

1.1 Introducción

En el capítulo actual se explican conceptos, características y elementos básicos acerca de los Almacenes de Datos y Mercados de Datos que permiten entender los principios de los mismos. Además se realiza una breve descripción de la metodología y herramientas utilizadas en la creación de los Mercados de Datos.

1.2 Almacenes de Datos

Los Almacenes de Datos tienen como objetivo organizar y permitir la consulta de la información con inmediatez. Estos son muy necesarios en el entorno empresarial para tomar mejores decisiones. La definición de Almacenes de Datos ha sido planteada por varios autores, estas definiciones enfatizan el uso de la información dentro de una empresa para apoyar a la toma de decisiones.

El concepto de Almacenes de Datos proviene de la palabra en inglés Data Warehouse, cuyo padre del término es William H. Inmon, el cual plantea "...un Almacén de Datos es una colección de datos orientado por temas, integrado, no volátil y variable en el tiempo que apoya las decisiones administrativas". (3)

Otro de los autores reconocidos en este ámbito es Ralph Kimball, expresa que un Almacén de Datos es "...una copia de datos transaccionales, específicamente estructurados para la consulta y el análisis". (4)

En la presente investigación se concluye que un Almacén de Datos es la unión de varios Mercados de Datos (subconjunto del Almacén) que contienen información y está orientado a un área específica de la empresa, obteniéndola de una o varias fuentes para su posterior análisis y que persistirá en el tiempo.

1.2.1 Características de los Almacenes de Datos

Como bien propone la definición de Inmon (3) los Almacenes de Datos constan de cuatro características principales.

Orientado por temas: el Almacén de Datos está orientado por las principales áreas o temáticas de la empresa. Los datos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales.

Integrado: el aspecto más importante del ambiente del Almacén de Datos es que la información encontrada en el interior está siempre integrada. La integración de datos se muestra de muchas maneras: en convenciones de nombres consistentes, en la medida uniforme de variables, en la codificación de estructuras consistentes, en atributos físicos de los datos consistentes, fuentes múltiples y otros.

No volátil: los datos almacenados no se modifican ni actualizan, sólo se añaden nuevos datos. Los datos de un Almacén existen para ser leídos, y no modificados, por tanto se carga una sola vez y siguen igual en lo adelante.

Variable en el tiempo: en los sistemas operacionales, los datos siempre reflejan el estado de la actividad del negocio en el presente. Por el contrario, la información en el Almacén sirve, entre otras cosas, para realizar análisis de tendencias. Por lo tanto, el Almacén se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones.

Ventajas y desventajas de los Almacenes de Datos

El uso de los Almacenes de Datos en empresas que manejan grandes volúmenes de información es vital ya que es una vía para tomar decisiones de forma clara, concisa y rápida. A continuación se proponen algunas ventajas y desventajas que poseen los Almacenes de Datos.

Poseen la ventaja de permitir el acceso rápido a una gran cantidad de datos en línea, debido a la capacidad que tienen para almacenar, estructurar y hacerlos disponibles. Además facilitan el funcionamiento de las aplicaciones de los sistemas de apoyo a la decisión tales como informes de tendencia e informes de excepción. Integran datos históricos sobre la actividad de la organización en un único repositorio, permitiendo analizar los datos del negocio desde la perspectiva de su evolución en el tiempo e identificar nuevas oportunidades, que posibiliten tomar decisiones estratégicas.

Como desventaja se encuentra que el ciclo de vida puede suponer altos costos, no suele ser estático y los costos de mantenimiento son elevados. Se tiene el riesgo de fracasar en la construcción del sistema, al subestimar los costes de captura y preparación de los datos, así como cambios frecuentes en los requisitos de los usuarios.

1.3 Mercado de Datos

Un Mercado de Datos, almacena la información de un área o departamento específico dentro del negocio. “Un Mercado de Datos es una solución que, compartiendo tecnología con el Almacén de Datos (pero con contenidos específicos, volumen de datos más limitado y un alcance histórico menor), permite dar soporte a una empresa pequeña, un departamento o área de negocio de una empresa grande”. (4)

Generalmente los Almacenes de Datos se subdividen en unidades lógicas más pequeñas dependiendo del subsistema de la entidad del que procedan o para el que sea necesario. Normalmente, un Mercado de Datos resuelve estudios a nivel de departamento. Un Mercado de Datos es una versión del Almacén de Datos. La diferencia principal es que la creación de un Mercado de Datos es específica para una necesidad de datos seleccionados, enfatizando el fácil acceso a una información relevante, es decir, los Mercados de Datos representan un subconjunto del Almacén de Datos, referentes a los requisitos de un departamento o área de negocio concreto con el objetivo de facilitar la construcción y utilización de un Almacén de Datos.

1.3.1 Características de los Mercados de Datos

- Según las necesidades de los usuarios el diseño del Mercado de Datos se realiza siguiendo una estructura consistente.
 - La información histórica que posee es mínima comparada con la información histórica que guardan los Almacenes de Datos.
 - Presentan mayor nivel de detalle, por lo que contienen el grado de granularidad necesaria.
 - A la hora de construirlo presenta costes adicionales en hardware, software y accesos de red.
 - Debido a que hay grupos de usuarios que solo acceden a un subconjunto preciso de datos, se hace más fácil el acceso a las herramientas de consulta y divide los datos para controlar mejores accesos.
- (4)

1.4 Arquitectura de un Almacén de Datos

La arquitectura de un Almacén de Datos viene determinada por su situación central como fuente de información para las herramientas de análisis. Su estructura básica incluye:

- Datos operacionales: fuente de datos para el componente de almacenamiento físico.

- Extracción de datos: selección sistemática de datos operacionales usados para poblar el componente de almacenamiento físico.
- Transformación de datos: Procesos para sumarizar y realizar otros cambios en los datos operacionales y para reunir los objetivos de orientación a temas e integración.
- Carga de datos: inserción de datos en el componente de almacenamiento físico mensualmente.
- Almacén: almacenamiento físico de datos.
- Herramientas de acceso: herramientas que proveen acceso a los datos.

1.5 Metodología de Desarrollo

Existen diversas metodologías que pretenden dar un acercamiento a una propuesta ideal para el desarrollo de un Almacén de Datos.

Entre las metodologías para la construcción de un Almacén de Datos se destacan la metodología descendente o top-down propuesta por Bill Inmon, como se muestra en la Figura 1 y la metodología ascendente o bottom-up propuesta por Ralph Kimball como se muestra en la Figura 2.

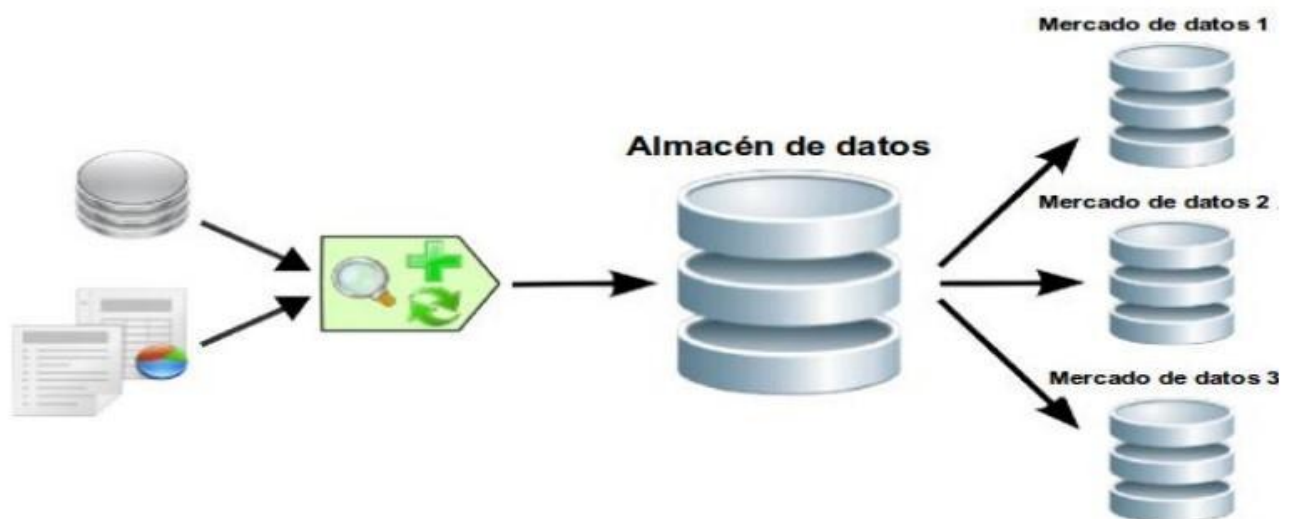


Figura 1. Enfoque descendente o top-down definido por Bill Inmon

La metodología de Kimball propone una arquitectura ascendente, este enfoque consiste en construir primeramente los Mercados de Datos de la entidad en cuestión, para luego unirlos y conformar el Almacén

de Datos. Por su parte Inmon define una metodología descendente a la hora de diseñar un Almacén de Datos, en este enfoque los mercados se crearán después de haber terminado el Almacén completo de la organización, por lo que es necesario conocer todos los aspectos de la empresa para poder realizar el mismo, ello trae consigo varios cambios a la hora de mejorar los procesos.

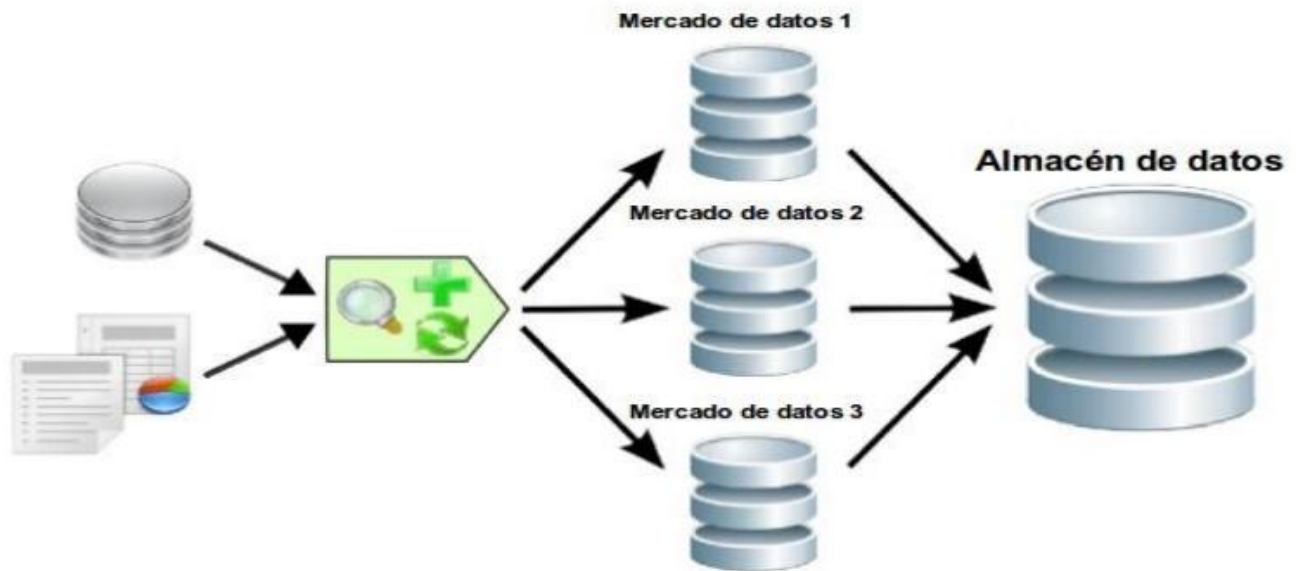


Figura 2. Enfoque ascendente o bottom-up definido por Ralph Kimball

Luego de realizar un estudio de las dos metodologías se puede concluir que la propuesta por Inmon es más compleja y requiere más tiempo para su implementación. Sin embargo, Kimball brinda una metodología de fácil comprensión y rápida de implementar por etapas, debido a estas razones y a que el proyecto SIGEF no ha terminado su implementación se decide utilizar la arquitectura para Almacenes de Datos propuesta por Kimball, para realizar tres Mercados de Datos que contribuyan con el análisis de los datos generados en los departamentos Verificación Fiscal, Quejas Peticiones y Denuncias, Dictámenes y Revisiones de Causas Penales del Sistema de Gestión de la Fiscalía.

1.6 Modelos de datos

Existen diversos tipos de implementaciones de la tecnología OLAP (On Line Analytic Processing), las que varían según el tipo de motor en el que se almacenan los datos. Los cuales se pueden clasificar de la siguiente manera:

ROLAP (Relational On Line Analytic Processing): son sistemas en los cuales los datos se encuentran almacenados en una Base de Datos relacional. Este tipo de organización física se implementa sobre tecnología relacional, pero disponen de algunas facilidades para mejorar el rendimiento.

MOLAP (OLAP Multidimensional): en estos sistemas se encuentran almacenados los datos en una estructura de datos multidimensional. De manera que la representación externa y la interna coincidan.

HOLAP (Hybrid On Line Analytic Processing): el procesamiento analítico híbrido en línea, constituye la unión entre MOLAP y ROLAP, combinando estas dos implementaciones para almacenar algunos datos en un motor relacional y otros en una Base de Datos multidimensional. (5)

Se decide que el tipo de almacenamiento será **MOLAP**, este tipo de modelado es más rápido que el **ROLAP** y es el que posee un mejor funcionamiento en los tiempos de respuesta a consultas. Un servidor MOLAP utiliza tecnología multidimensional donde los datos están almacenados en matrices. Las operaciones de consultas están implementadas directamente sobre estas matrices, la información se almacena a través de tablas de hechos y tablas de dimensiones.

1.7 Base de Datos multidimensional

Las Bases de Datos multidimensional utilizan dos tipos de tablas en sus Bases de Datos, estas son las tablas de hechos y las tablas de dimensiones. Además tienen tres variantes posibles de modelado: esquema en estrella, esquema constelación y el esquema copo de nieve.

Tablas de hechos: Las tablas de hechos son donde las mediciones numéricas del negocio son almacenadas. Cada una de las mediciones es tomada como la intersección de todas las dimensiones. (6)

Habitualmente cada tabla de hecho define un Mercado de Datos determinado, debido a que en ellas se almacena la información perteneciente al tema en cuestión, ejemplo: cuadros, reservas, evaluaciones. La principal condición que deben cumplir las tablas de hechos es que el hecho debe almacenarse de tal forma

que su valor sea numérico y a su vez sea aditivo para así poder realizar cálculos sobre él, ya sea por ciento, sumas, entre otros.

Tablas de dimensiones: Las tablas dimensionales son aquellas donde las descripciones textuales de las dimensiones del negocio son almacenadas. Cada una de las descripciones textuales ayuda a describir un miembro de la dimensión respectiva. (6)

Las tablas de dimensiones concretan como están los datos introducidos lógicamente y proporcionan el medio para analizar el negocio. Representan los ejes del cubo, y los aspectos de interés, mediante los cuales el usuario podrá filtrar y manipular la información almacenada en la tabla de hechos.

Dentro de las tablas podemos encontrar:

Jerarquía: Es un conjunto de atributos descriptivos que se rigen por un orden preestablecido. Una jerarquía implica una organización de niveles dentro de una dimensión, con cada nivel representando el total agregado de los datos del nivel inferior.

Granularidad: Es el nivel más bajo de información que será almacenado en la tabla de hechos e indica el grado de detalle asociado a un hecho particular. La granularidad depende directamente del número de dimensiones que se asocian a la tabla de hechos.

Los datos dentro de la tabla dimensión, que proveen información del negocio o que describen alguna de sus características, son llamados datos de referencia.

Entonces, se puede afirmar que una tabla de dimensión posee una clave primaria y uno o más datos de referencia.

1.7.1 Variantes de modelado

Esquema en estrella: posee una gran tabla central (tabla de hechos) y un conjunto de pequeñas tablas acompañantes (tablas dimensiones) presentadas en un modelo radial alrededor de la tabla central como se muestra en la Figura 3.

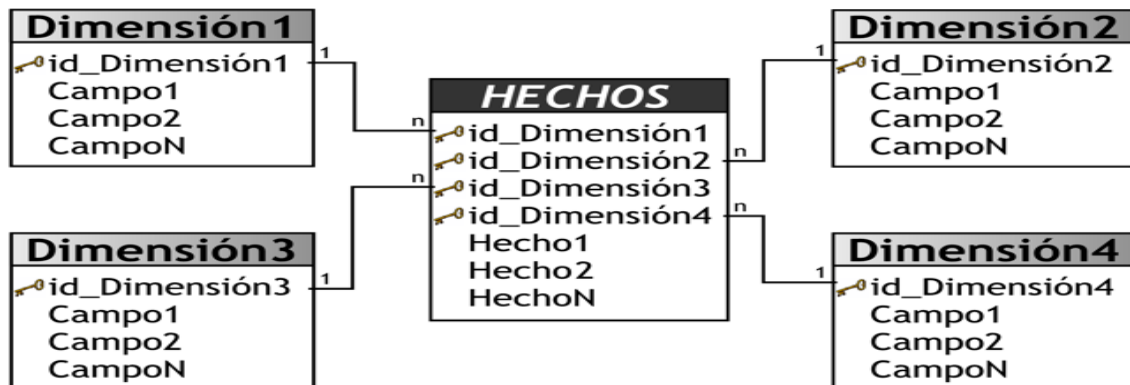


Figura 3. Esquema en estrella

Esquema constelación: este modelo está formado por una tabla de hechos principal y por una o más tablas de hechos auxiliares las cuales pueden ser sumalizaciones de la principal. Dichas tablas yacen en el centro del modelo y están relacionadas con sus respectivas tablas de dimensiones como lo muestra la Figura 4. (7)

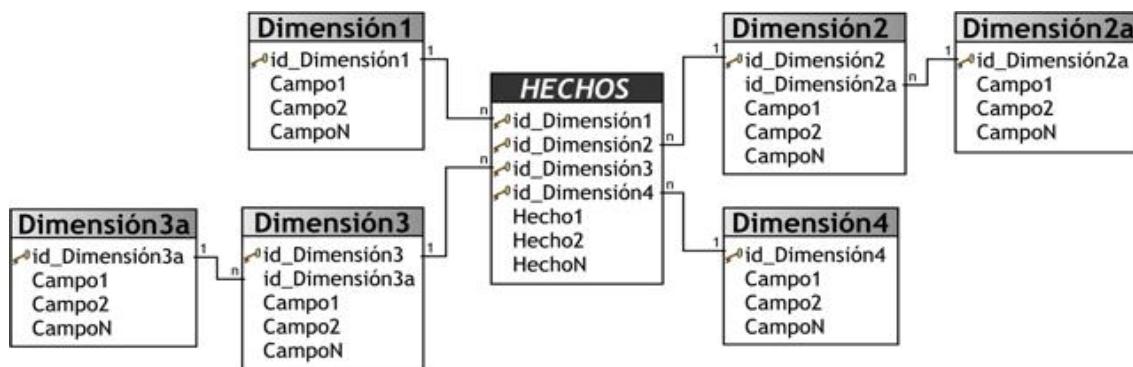


Figura 4. Esquema en copo de nieve.

Esquema copo de nieve: este esquema representa una extensión del modelo en estrella cuando las dimensiones se organizan en jerarquías de dimensiones como se muestra en la Figura 5. (7)

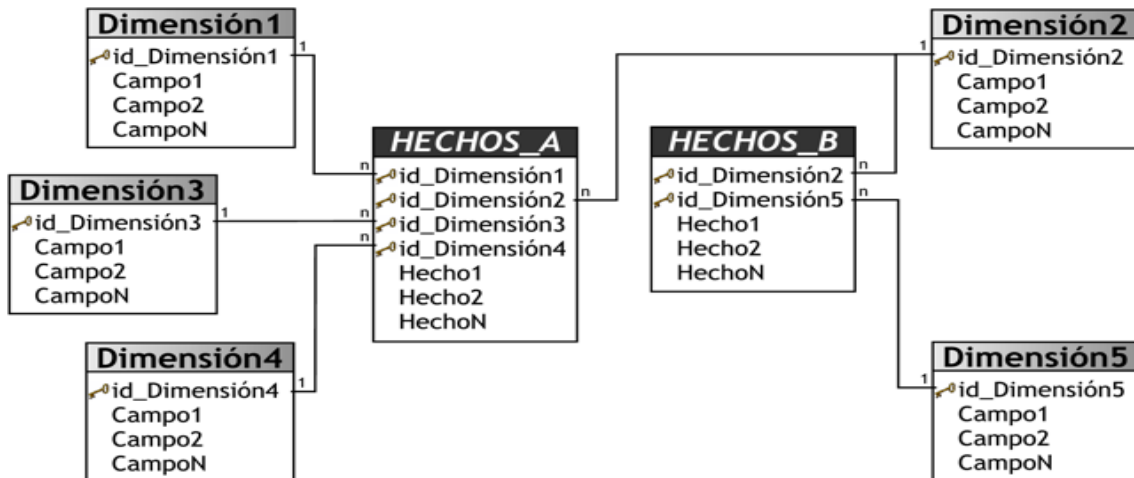


Figura 5. Esquema de constelación de hechos.

Entre los tres esquemas dimensionales especificados anteriormente, se estableció hacer uso de los esquemas constelación de hechos para el modelado de los datos de los Mercados de Datos en general. Se selecciona porque está más acorde a las características del negocio debido a que cada subsistema está compuesto por más de un proceso.

1.8 Extracción, Transformación y Carga de los datos

La extracción, transformación y carga de los datos (ETL por sus siglas en inglés) es el proceso que organiza el flujo de los datos entre diferentes sistemas en una organización y aporta los métodos y herramientas necesarias para mover datos desde múltiples fuentes a un Almacén de Datos.

La idea es que una aplicación ETL lea los datos primarios de unas Bases de Datos de sistemas principales, realice transformación, validación, el proceso cualitativo, filtración y al final escriba los datos en el Almacén, dejándolos en ese momento disponible para analizar por los usuarios.

La **extracción** se basa en la acción de obtener la información deseada a partir de los datos almacenados en fuentes externas, la **transformación** es la operación realizada sobre los datos para que puedan ser cargados en el Almacén de Datos y la **carga** consiste en almacenar los datos en la Base de Datos final.

Si no se realiza un correcto proceso de ETL se pueden obtener datos incorrectos lo que afectaría el proceso de toma de decisiones, es por esto que este proceso constituye aproximadamente un 70% del trabajo de la construcción de un Mercado de Datos. (8)

1.9 Herramienta de modelado

Las herramientas de modelado se utilizan para representar los elementos claves del proceso de manera que sea posible alcanzar una mejor comprensión del mismo. (9) Las herramientas de Ingeniería Asistida por Computadora (Computer Aided Software Engineering, CASE por sus siglas en inglés), tienen como objetivo incrementar la productividad y calidad de los productos de software, mejorar la planificación del proyecto, así como reducir el tiempo y costo de su desarrollo. Existen varias herramientas CASE, entre las que se encuentran: ER Estudio, Rational Rose y Visual Paradigm.

1.9.1 Visual Paradigm 8.0

Visual Paradigm es una herramienta CASE de diseño que soporta los principales estándares de la industria donde se encuentra el Lenguaje de Modelado Unificado (Unified Modeling Language, UML por sus siglas en inglés). Ofrece un conjunto completo de herramientas para el desarrollo del software, necesarias para la captura de requisitos, software de planificación, planificación de controles, modelado de clases y de datos. (10)

En su versión 8.0 brinda la posibilidad de modelar numerosos tipos de diagramas de clases, permite la compatibilidad entre ediciones, la generación de código desde diagramas y la documentación asociada a cada etapa del proceso de desarrollo. Además presenta licencia gratuita y comercial, por lo que todas las personas pueden trabajar con dicha herramienta.

Se decidió utilizar Visual Paradigm 8.0 para UML porque es una herramienta CASE profesional que soporta todo el ciclo de vida del desarrollo de software. Además, brinda la posibilidad de modelar todo tipo de diagramas de clases, admite la compatibilidad entre ediciones, la documentación asociada a cada etapa del proceso de desarrollo y generar script para diferentes sistemas gestores de Base de Datos. Los desarrolladores lo utilizan para facilitar el modelado simultáneo, almacenar los archivos de proyectos y hacer un seguimiento de los cambios.

1.10 Sistema Gestor de Base de Datos

Un Sistema Gestor de Bases de Datos (DataBase Management System, en inglés DBMS) es un software que permite definir Bases de Datos, estructurar los datos que serán almacenados y la búsqueda de los mismos, ya sea de forma interactiva o a través de otras tecnologías. Es una herramienta efectiva que permite

a varios usuarios acceder a la información al mismo tiempo. Brindan facilidades y un grupo de funciones con el objetivo de garantizar la confidencialidad, calidad, seguridad e integridad de los datos que contienen, así como el acceso fácil y relativamente rápido de los mismos. Dentro de los SGBD (Sistemas Gestores de Bases de Datos), se pueden encontrar el PostgreSQL, Oracle, Sybase, MySQL, entre otros.

1.10.1 PostgreSQL 9.1

PostgreSQL es un SGBD objeto-relacional, desarrollado en la Universidad de California, en el Departamento de Ciencias de la Computación de Berkeley. Brinda un control de concurrencia multi-versión (Multiversion concurrency control, MVCC por sus siglas en inglés) que permite trabajar con grandes volúmenes de datos. La versión 9.1 es estable y segura, además de proporcionar la elaboración de consultas en SQL de forma gráfica.

Es el SGBD de código abierto más potente en el mercado y en sus últimas versiones ha mejorado varios factores que comparado con otros SGDB exponía ciertas desventajas. PostgreSQL utiliza un modelo cliente/servidor y usa multiprocesos en vez de multihilos para garantizar la estabilidad del sistema. Un fallo en uno de los procesos no afectará el resto y el sistema continuará funcionando. (11)

PostgreSQL 9.1 posee características técnicas que lo hacen una de las Bases de Datos más potentes y robustas del mercado. Las características que más se han tenido en cuenta durante su desarrollo son la estabilidad, potencia, robustez, facilidad de administración e implementación de estándares. PostgreSQL funciona muy bien con grandes cantidades de datos y una alta concurrencia de usuarios accediendo a la vez al sistema. (12) Es por esto que se utilizará PostgreSQL en su versión 9.1, además de haber sido este el seleccionado por el equipo de desarrollo para la creación del Mercado de Datos.

1.10.2 Administrador de Base de Datos. PgAdmin III 1.14.0

Un administrador de Base de Datos (DataBase Administrator, DBA por sus siglas en inglés) es fundamental para el desarrollo de una investigación de este tipo, ya que son los responsables de la integridad y disponibilidad de los datos. Como parte de su función, crean y configuran Bases de Datos relacionales. También se encargan de llevar a cabo el diseño de la distribución de los mismos y las soluciones de almacenamiento, el despliegue y monitorización de servidores de Bases de Datos. Algo muy importante y

que a todas las instituciones les interesa es que garantizan la seguridad de las Bases de Datos, incluyendo backups y recuperación de desastres.

PgAdmin III es una aplicación gráfica para administrar el SGBD PostgreSQL, siendo la más completa y popular con licencia Open Source¹. (11) Este software fue diseñado para responder a las necesidades de todos los usuarios, desde la escritura de simples consultas SQL a la elaboración de Bases de Datos complejas. La interfaz gráfica es compatible con todas las características de PostgreSQL y facilita la administración. La aplicación también incluye un editor de la sintaxis SQL. PgAdmin III es capaz de gestionar versiones a partir de la PostgreSQL 7.3 ejecutándose en cualquier plataforma.

1.11 Herramientas para la Extracción, Transformación y Carga de datos

El proceso de ETL es uno de los más importantes en la construcción de un Mercado de Datos, debido a que es necesario que la información contenida represente la realidad del negocio, sea confiable y esté disponible en el momento que los usuarios y la organización la necesiten. Por la importancia que tienen estos procesos, es preciso contar con herramientas que permitan reducir tiempo y costos. Una de las herramientas que nos brindan esta posibilidad es Pentaho Data Integration o por sus siglas en inglés PDI. (13)

1.11.1 Pentaho Data Integration 4.2.1

Pentaho Data Integration (PDI, también conocida como Kettle) es una de las soluciones más extendidas y mejor valoradas en el mercado, que reúne un conjunto de componentes que permiten modelar y ejecutar transformaciones sobre flujos de datos. Posee capacidades de integración de datos, entorno de diseño gráfico intuitivo rico y una arquitectura altamente escalable, proporciona la solución ideal para cualquier tipo de integración de datos, análisis de negocio o proyectos con grandes capacidades de datos. (14)

Es una herramienta multiplataforma, lo que permite ejecutarla en cualquier sistema operativo. Está basada en dos tipos de objetos; las transformaciones que contienen una colección de pasos en un proceso de ETL y los trabajos que poseen una colección de transformaciones y/o trabajos. Además brinda soporte para metadatos, así como funciones que permiten operar con los campos en el flujo de datos, renombrando, calculando campos en función de otros, correlacionando valores y realizando búsquedas auxiliares en

¹ Open Source es el término con el que se conoce al software distribuido y desarrollado libremente.

Bases de Datos. Ofrece soporte para operaciones de dimensiones lentamente cambiantes, permite ejecutar código Java Script dentro de las transformaciones e incorpora un evaluador de expresiones regulares.

Además esta herramienta es fácil de usar, brinda la posibilidad de copiar y leer del mismo fichero en paralelo, permitiendo maximizar la capacidad de entrada/salida en el entorno ETL. Añade un debugger² integrado diseñado para mejorar la productividad del desarrollador, ya que se pueden agregar puntos de ruptura condicionales en la ejecución de las transformaciones, dando la posibilidad de pausar y resumir la ejecución de la transformación, así como especificar el número de filas que se van a usar en las ejecuciones de prueba. Además, se pueden añadir registros personalizados. Cuenta con una gran comunidad de usuarios.

Presenta algunas desventajas, ya que no cuenta con un componente de calidad de datos, no automatiza el proceso de separación y redistribución de datos para el procesamiento paralelo. Además para realizar búsquedas de mayores volúmenes necesita utilizar una Base de Datos de búsqueda donde se ejecutan un gran número de sentencias SQL (structured query language) que frenan el rendimiento de la ETL.

En esta investigación el enfoque está dirigido a las herramientas de código abierto, pues aporta la posibilidad de manejar directamente el código fuente, modificarlo o adaptarlo a las necesidades individuales.

1.11.2 Mondrian Schema WorkBench

WorkBench en su versión 3.2.0 es una herramienta para el desarrollo del esquema del modelo estrella en XML Work Bench desarrollada en Java. Esta herramienta es una interfaz de diseño que permite crear y probar esquemas de cubos OLAP visualmente. Este programa entrega todas las facilidades para poder realizar el modelo lógico del cubo OLAP al cual se le realizarán las consultas. Los modelos de esquemas XML de metadatos se crean en una estructura específica utilizada por el motor de Mondrian. La estructura de estos modelos se pueden considerar de forma de cubos, que utilizan hechos existentes y tablas de dimensiones que se encuentran en el gestor de Base de Datos. (15)

1.11.3 Pentaho BI server 4.8.0

² Es un programa usado para probar y depurar (eliminar los errores) de otros programas.

La aplicación más conocida de la Plataforma de Bi es la Pentaho BI Server que funciona como una red basada en sistema de gestión de informe, el servidor de integración de aplicaciones y un motor de flujo de trabajo ligero (secuencias de acción.) Está diseñado para integrarse fácilmente en cualquier proceso de negocio. (16)

1.12 Herramientas para las pruebas.

1.12.1 Apache JMeter 2.9.

Es un software de código abierto realizado en Java, puede ser utilizado para probar el rendimiento tanto en recursos estáticos y dinámicos (archivos, Servlets, scripts de Perl, Java Objects, Bases de Datos y consultas, servidores FTP y mucho más). Se puede utilizar para simular una carga pesada en el servidor, de red o un objeto para probar su resistencia o para analizar el rendimiento general bajo diferentes tipos de carga. Se puede utilizar para hacer un análisis gráfico de rendimiento o para probar su servidor / script / comportamiento del objeto bajo carga pesada concurrentes. (17)

1.12.2 Datanamic Data Generator for PostgreSQL 5.0.

Datanamic Data Generator es una herramienta para la generación de datos de prueba a tablas de Bases de Datos en PostgreSQL. Permite definir tablas y campos para la generación de datos, establecer rangos de valores, generar campos char por la máscara, obtener listas de valores de las consultas SQL y muchas otras características para generar datos de prueba de forma sencilla y de manera directa. También proporciona aplicación de consola, lo que le permite generar datos en un solo toque mediante el uso de plantillas de generación. Las principales características que posee son. (11)

- Fácil de usar interfaz de asistente
- Seis idiomas disponibles: Inglés, francés, alemán, italiano, ruso y español.
- Generación de datos a varias tablas de Bases de Datos diferentes en un host.
- Todos los tipos de datos PostgreSQL, incluyendo Array, direcciones de red y tipos geométricos.
- Los diferentes tipos de generación para cada campo, incluyendo lista, la generación incremental de datos al azar y mucho más.
- Capacidad para utilizar los resultados de consultas SQL como lista de valores para la generación de datos.

- El control automático sobre la integridad referencial para las tablas vinculadas generación de datos.
- Gran variedad de parámetros de generación para cada tipo de campo.
- Capacidad para establecer los valores NULL para cierto porcentaje de los casos.
- Posibilidad de guardar todos los parámetros de generación, creado en la sesión del asistente actual.
- Utilidad de línea de comandos para generar datos utilizando el archivo de plantilla.

1.13 Conclusiones del capítulo

Una vez concluido el estudio de las diferentes tecnologías para dar solución al problema planteado, se puede confirmar que los Almacenes de Datos constituyen un elemento primordial en la gestión de grandes volúmenes de información, auxiliándose en los Mercados de Datos para informaciones clasificadas específicamente. Se abordaron los principales elementos que componen los Almacenes de Datos. Así como las herramientas y tecnologías a utilizar para el desarrollo de los Mercados de Datos.

Para la implementación de la solución se concluye que:

- La Propuesta de Metodología para el Desarrollo de Almacenes de Datos permite guiar el proceso de desarrollo del Mercado de Datos por las diferentes etapas del ciclo de vida.
- Las herramientas que ayudarán a construir el subsistema de almacenamiento son: Visual Paradigm en su versión 8.0 que posibilita la generación de los diagramas necesarios para modelar el funcionamiento del sistema propuesto. Así como el SGBD PostgreSQL 9.1 y el administrador de Base de Datos PgAdmin III 1.14.0.
- Se decide emplear el sistema de almacenamiento MOLAP, ya que este tipo de modelado es más rápido que el ROLAP.
- Para la implementación del subsistema de integración se decide utilizar el Pentaho Data Integration 4.2.1 para extraer, transformar y cargar los datos.

2 Capítulo 2. Diseño e implementación de los Mercados de Datos.

2.1 Introducción

En este capítulo se realiza el análisis de los requerimientos donde se identificarán las necesidades de información de los usuarios, luego se realiza un análisis de las fuentes de datos existentes donde se determina el estado general de las mismas. Se construye la matriz dimensional, el modelo de datos donde se determinan: las dimensiones, los hechos y las medidas además se realizarán las uniones entre las mismas para conformar el modelado de los Mercados de Datos. Se realiza la implementación de los Mercados de Datos para los subsistemas CLEP, PDC y VF, este proceso se efectúa sobre las tablas de los modelos de datos.

2.2 Análisis de las necesidades de información

La primera etapa del desarrollo de la solución comienza identificando las necesidades de información de los usuarios, debido a que es importante conocer cuál es la información que necesitan y como es que la desean. A continuación se muestran algunas de las necesidades de información más significativas que se identificaron después de realizar un análisis de los reportes generados en los departamentos Verificación fiscal, Quejas, peticiones y denuncias, Dictámenes y Revisiones de causas penales las restantes se pueden consultar en el Anexo 1:

Para el subsistema Verificación Fiscal:

Total de acciones de control planificadas.

- cantidad de Verificaciones Fiscales.
- cantidad de Investigaciones.

Cantidad de acciones planificadas sectorializadas.

- cantidad de Verificaciones Fiscales.
- cantidad de Investigaciones.

Total de reinspecciones planificadas.

- cantidad asociadas a verificaciones fiscales.

- cantidad asociadas a investigaciones fiscales.
- cantidad de reinspecciones sectorializadas.
- cantidad asociadas a verificaciones fiscales.
- cantidad asociadas a investigaciones fiscales.

Total de acciones de control iniciadas.

- cantidad de Verificaciones Fiscales.
- cantidad de Investigaciones.

Cantidad de acciones iniciadas sectorializadas.

- cantidad de Verificaciones Fiscales.
- cantidad de Investigaciones.

Total de acciones de control concluidas.

- cantidad de Verificaciones Fiscales.
- cantidad de Investigaciones.

Total de acciones de control concluidas en el término.

- cantidad de Verificaciones Fiscales.
- cantidad de Investigaciones.

Para el subsistema CLEP en el departamento Quejas, Peticiones y Denuncias:

Total de quejas, peticiones y denuncias recibidas

Total de quejas recibidas

Total de quejas radicadas.

Total de quejas trasladadas.

Total de quejas Tramitadas.

Total de quejas archivadas.

Cantidad de quejas tramitadas en término.

Cantidad de quejas resueltas fuera del término.

Total de denuncias recibidas.

Total de denuncias trasladadas.

Total trasladadas a la PNR.

Total trasladadas al Órgano de Instrucción.

Total trasladadas a la Fiscalía Militar.

Para el subsistema PDC en el departamento Revisiones de causas penales:

Total de solicitudes de revisión penales recibidas en la fiscalía.

Cantidad de solicitudes de revisión penales recibidas por género.

Cantidad de solicitudes de revisión penales recibidas por género femenino.

Cantidad de solicitudes de revisión penales recibidas por género masculino.

Cantidad de solicitudes trasladadas.

Cantidad de solicitudes archivadas.

Cantidad solicitudes trasladadas por género femenino.

Cantidad solicitudes trasladadas por género masculino.

Cantidad solicitudes trasladadas por género.

Cantidad de solicitudes de pruebas.

Para el subsistema CLEP en el departamento Dictámenes:

Total de dictámenes favorables del Fiscal por concepto de salida excepcional del territorio nacional presentado por la DEP.

Total de dictámenes desfavorables del Fiscal por concepto de salida excepcional del territorio nacional presentado por la DEP.

Total de autos recibidos de las sustitutivas presentadas por la DEP.

Total de autos recibidos de las sustitutivas presentadas por la PNR.

Total de recursos de súplica interpuestos.

Total de recursos de súplica declarados con lugar.

Total de recursos de súplica declarados sin lugar.

Total de libertades condicionales presentadas.

Total de libertades condicionales dictaminadas favorables por el Fiscal.

Total de libertades condicionales dictaminadas desfavorables por el Fiscal.

Total de libertades condicionales presentadas al tribunal.

Total de libertades condicionales aprobadas por el tribunal.

2.3 Análisis del estado de las fuentes de datos

Para gestionar los datos correspondientes al sistema SIGEF se cuenta con una Base de Datos relacional y como gestor de Base de Datos se utiliza PostgreSQL en la versión 9.1.

Como medida organizativa y técnica tomada para asegurar el funcionamiento del sistema SIGEF, los servidores deben estar disponibles las 24 horas del día y por consiguiente la Base de Datos se debe encontrar en plena disponibilidad. El sistema SIGEF cuenta con una funcionalidad para la réplica de datos, por lo que los servidores de los diferentes niveles se encuentran actualizados.

El departamento Verificación fiscal tiene un total de 97 tablas de las cuales 20 son tablas nomencladoras y 77 son tablas de datos, el departamento Quejas, peticiones y denuncias tiene un total de 19 tablas de las cuales 6 son tablas nomencladoras y 13 son tablas de datos, el departamento Dictámenes tiene un total de 24 tablas de las cuales 13 son tablas nomencladoras y 11 son tablas de datos y el departamento Revisiones de casusas penales tiene un total de 39 tablas de las cuales 11 son tablas nomencladoras y 28 son tablas de datos. Se pueden señalar como tablas relevantes para los Mercados de Datos las tablas. Además fue necesaria la utilización de algunas tablas del subsistema Base entre estas se encuentran: dfiscalia, ddocumento y dproceso.

2.4 Metodología de Kimball

Seguidamente se le da cumplimiento a los pasos propuestos por la metodología de Kimball para el desarrollo de los Mercados de Datos de los subsistemas CLEP, PDC y VF.

2.4.1 Planificación

Alcance: Los Mercados de Datos serán realizados para los departamentos: Revisiones de causas penales del subsistema PDC, Quejas, peticiones y denuncias del subsistema CLEP y Verificación fiscal del SIGEF.

2.4.2 Análisis de requerimientos

En esta fase se realiza una reunión con los analistas del SIGEF para determinar las necesidades de información relevantes para las Fiscalías referentes a los subsistemas CLEP, PDC y VF.

De esta reunión se seleccionaron los temas analíticos y sus requerimientos como se muestran a continuación en las Tablas 1, 2 y 3.

Tema analítico	Análisis o requerimiento inferido o pedido.	Proceso de negocio de soporte.	Comentario
Dictámenes	Análisis de los dictámenes.	Dictámenes	Por concepto de libertad condicional, por concepto de libertad condicional extraordinaria.
Quejas, peticiones y denuncias	Análisis de los escritos recibidos por la fiscalía.	Registro de escrito	Por tema, por traslado, por radicación.
	Análisis de los escritos tramitados por la fiscalía.	Rollo de quejas	Por término, por tramitación, por traslado, por archivo.

Tabla 1. Temas analíticos. Subsistema CLEP.

Tema analítico	Análisis o requerimiento inferido o pedido.	Proceso de negocio de soporte.	Comentario
Verificación fiscal	Análisis de las acciones de control.	Acciones	Por verificaciones fiscales, por investigaciones fiscales, por reinspecciones, por sectorializadas.
	Análisis de los hechos delictivos y violaciones.	Hechos	Por verificaciones fiscales, por investigaciones fiscales, por reinspecciones, por sectorializadas.

Tabla 2. Temas analíticos. Subsistema VF.

Tema analítico	Análisis o requerimiento inferido o pedido.	Proceso de negocio de soporte.	Comentario
Revisiones de causas penales	Análisis de las solicitudes recibidas por la fiscalía.	Solicitudes recibidas	Por período, por solicitudes recibidas, por solicitudes trasladadas.

	Análisis de las solicitudes de pruebas.	Rollo de tramitación	Por período, por revisiones, por escrito promoción, por dictamen denegatorio.
		Rollo de FGR	Por tipo de prueba.

Tabla 3. Temas analíticos. Subsistema PDC.

2.4.3 Matriz bus o matriz dimensional

La matriz dimensional especifica la relación entre los hechos y las dimensiones, donde las columnas contienen las tablas de hechos y las filas las dimensiones, la intersección de una fila con una columna especifica si hay relación entre una tabla de hechos y una dimensión.

Matriz dimensional para el subsistema CLEP

Dimensiones	Hechos		
	H1	H2	H3
Tema	X		
Traslado órgano	X		
Decisión depurar	X		
Prórroga		X	
Término		X	X
Fallo del tribunal		X	X
Procedencia			X
Tipo subsidiaria			X

Tipo dictamen			X
Decisión dictamen			X
Tiempo	X	X	X
Tipo resolución			X
Persona	X		X
Instancia	X	X	X
Tipo escrito	X		X
Decisión tramitar		X	

Leyenda:

H1: Hecho registro escrito.

H2: Hecho rollo de quejas.

H3: Hecho dictamen.

Matriz dimensional para el subsistema VF

Dimensiones	Hechos	
	H1	H2
Aprobación prórroga	X	
Término	X	X
Fallo del tribunal	X	X
Tiempo	X	X

Instancia	X	X
Tipo de responsable		X
Acciones concluidas	X	
Violaciones		X
Medidas disciplinarias		X
Categoría		X
Acciones iniciadas	X	

Leyenda:

H1: Hecho acción.

H2: Hecho hechos.

Matriz dimensional para el subsistema PDC

Dimensiones	Hechos		
	H1	H2	H3
Decisión depurar	X	X	
Tipo de prueba			X
Fallo del tribunal		X	
Tiempo	X	X	X
Persona	X	X	
Instancia	X	X	X
Disposición fiscal		X	

Traslado penal	X		
Revisión		X	

Leyenda:

H1: Hecho solicitudes recibidas.

H2: Hecho rollo de tramitación.

H3: Hecho rollo de FGR.

2.4.4 Modelado Dimensional

Realizar las matrices dimensionales permite conocer que cada uno de los hechos definidos para los modelos de datos, comparten dimensiones, sin embargo no existen dos o más hechos que se relacionen con exactamente las mismas dimensiones. Esto indica que pudo verificarse la inexistencia de solapamiento entre hechos.

2.4.5 Elegir el proceso de negocio

Las áreas de proceso a modelar serían del subsistema PDC el departamento: Revisiones de causas Penales, del subsistema CLEP los departamentos Quejas, Peticiones y Denuncias y Dictámenes, y el subsistema Verificación Fiscal.

2.4.6 Establecer el nivel de granularidad

En este punto se muestra una tabla con la descripción de los datos asociados a las dimensiones, donde se explica el tipo de dato y/o los posibles valores que pueden tomar estos. En las Tablas 4, 5 y 6 se muestra la granularidad de las dimensiones identificadas para cada Mercado de Datos.

Nombre de la dimensión.	Nombre del dato	Tipo de dato	Valor nulo	Descripción

tiempo	id_tiempo	Numeric(10)	No	Identificador del tiempo.
	Mes	Numeric(10)	No	Mes en el que se crea el proceso.
	Dia	Numeric(10)	No	Día en el que se crea el proceso.
	Anno	Numeric(10)	No	Año en el que se crea el proceso.
	trimestre	Numeric(10)	No	Trimestre en el que se crea el proceso.
termino	id_termino	Numeric(10)	No	Identificador del término de tiempo.
	cant_dias	Numeric(10)	No	Cantidad de días en los que se debe concluir una acción determinada.
Instancia	descripcion	Varchar(255)	No	Instancia en que se analizan los datos (FGR, Provincial, Municipal).
	id_fiscalia	Numeric(19,0)	no	Identificador de la instancia.
Persona	id_persona	Numeric(19,0)	No	Identificador de la persona.
	Sexo	Varchar(255)	No	Sexo de la persona.

	Raza	Varchar(255)	No	Raza de la persona.
decision_depurar	id_decision_ado p	Numeric(19,0)	No	Identificador de la decisión de la depuración.
	Descripcion	Varchar(255)	No	Decisión del fiscal de la depuración.
fallo_tribunal	id_fallo	Numeric(19,0)	No	Identificador del fallo.
	Descripcion	Varchar(255)	No	Descripción del fallo del tribunal, si la respuesta es con lugar o no.
tipo_escrito	id_tipo_escrito	Numeric(19,0)	No	Identificador del tipo de escrito.
	Descripcion	Varchar(255)	No	Tipo de escrito recibido en la fiscalía.
traslado_organo	id_organo	Numeric(19,0)	No	Identificador de los órganos de traslado.
	Órgano	Varchar(255)	No	Órgano al que se traslada.
Tema	id_tema	Numeric(19,0)	No	Identificador del tema a que pertenecen los escritos.
	Descripcion	Varchar(255)	No	Tema a los que pertenecen los escritos.

decision_dictamen	id_decision_dictamen	Numeric(19,0)	No	Identificador de la decisión del dictamen.
	Descripcion	Varchar(255)	No	Contiene si esa decisión es favorable o desfavorable.
Prorroga	id_documento	Numeric(19,0)	No	Identificador de la prórroga.
	Decisión	Varchar(255)	No	Decisión del fiscal de otorgar o no la prórroga.
Procedencia	id_procedencia	Numeric(19,0)	No	Identificador de la procedencia.
	Descripcion	Varchar(255)	No	Contiene la procedencia.
tipo_subsidiaria	id_tipo_subsidiaria	Numeric(19,0)	No	Identificador del tipo de subsidiaria.
	Descripcion		No	Descripción del tipo de subsidiaria.
tipo_dictamen	id_tipo_dictamen	Numeric(19,0)	No	Identificador del tipo de dictamen.
	Descripcion		No	Descripción del tipo de dictamen.
decision_tramitar	id_decision_tramitar	Numeric(19,0)	No	Identificador de la decisión tramitar.
	discriminante		No	Decisión de la tramitación.

Tabla 4. Nivel de granularidad para el subsistema CLEP.

Nombre de la dimensión	Nombre del dato	Tipo de dato	Valor nulo	Descripción
Tiempo	id_tiempo	Numeric(10)	No	Identificador del tiempo.
	Mes	Numeric(10)	No	Mes en el que se crea un proceso.
	Dia	Numeric(10)	No	Día en el que se crea un proceso.
	Anno	Numeric(10)	No	Año en el que se crea un proceso.
	Trimestre	Numeric(10)	No	Trimestre en el que se crea un proceso.
Instancia	descripcion	Varchar(255)	No	Instancia en que se analizan los datos (FGR, Provincial, Municipal).
	id_fiscalia	Numeric(19,0)	no	Identificador de la instancia.
tipo_prueba	id_tipo_prueba	Numeric(19,0)	No	Identificador del tipo de prueba.
	descripcion	Varchar(255)	No	Descripción de los tipos de pruebas realizados.

Persona	id_persona	Numeric(19,0)	No	Identificador de la persona.
	Sexo	Varchar(255)	No	Sexo de la persona.
decision_adoptada	id_decision_adoptada	Numeric(19,0)	No	Identificador de la decisión adoptada por el fiscal.
	descripcion	Varchar(255)	No	Decisión adoptada por el fiscal.
traslado_penal	id_traslado	Numeric(19,0)	No	Identificador del traslado penal.
	descripcion	Varchar(255)	No	Tipo de traslado penal.
disposicion_fiscal	id_disposicion	Numeric(19,0)	No	Identificador de la disposición fiscal.
	decision_fiscal	Varchar(255)	No	Decisión del fiscal.
fallo_tribunal	id_fallo	Numeric(19,0)	No	Identificador del fallo.
	Descripcion	Varchar(255)	No	Descripción del fallo del tribunal, si la respuesta es con lugar o no.
Revision	id_revision	Numeric(19,0)	No	Identificador de la revisión.
	Descripcion	Varchar(255)	No	Descripción del tipo de revisión.

Tabla 5. Nivel de granularidad para el subsistema PDC.

Nombre de la dimensión	Nombre del dato	Tipo de dato	Valor nulo	Descripción
Tiempo	id_tiempo	Numeric(10)	No	Identificador del tiempo.
	Mes	Numeric(10)	No	Mes en el que se crea un proceso.
	Dia	Numeric(10)	No	Día en el que se crea un proceso.
	Anno	Numeric(10)	No	Año en el que se crea un proceso.
	Trimestre	Numeric(10)	No	Trimestre en el que se crea un proceso.
Instancia	Descripcion	Varchar(255)	No	Instancia en que se analizan los datos (FGR, Provincial, Municipal).
	id_fiscalia	Numeric(19,0)	no	Identificador de la instancia.
acciones_iniciadas	id_documento	Numeric(19,0)	No	Identificador de la acción iniciada.
	descripcion	Varchar(255)	No	Descripción de las acciones iniciadas.

tipo_responsable	id_tipo_responsable	Numeric(19,0)	No	Identificador del tipo de responsable.
	Descripcion	Varchar(255)	No	Descripción del tipo de responsable.
accion_concluida	id_documento	Numeric(19,0)	No	Identificador de la acción concluida.
	Descripcion	Varchar(255)	No	Descripción de la acción concluida.
fallo_tribunal	id_fallo	Numeric(19,0)	No	Identificador del fallo.
	Descripcion	Varchar(255)	No	Descripción del fallo del tribunal, si la respuesta es con lugar o no.
medidas_disciplinarias	id_med_solicita da	Numeric(19,0)	No	Identificador de la medida solicitada.
	Disciplinaria	Varchar(255)	No	Si la medida es o no disciplinaria.
Termino	id_termino	Numeric(19,0)	No	Identificador del término de tiempo.
	cant_dias	Numeric(19,0)	No	Cantidad de días en que se debe realizar una acción.

Violaciones	id_violaciones	Numeric(19,0)	No	Identificador de las violaciones detectadas.
	causas_condiciones	Varchar(255)	No	Causas de la violación.
aprobacion_prorroga	id_documento	Numeric()	No	Identificador de la prórroga.
	Decisión	Varchar(255)	No	Decisión del fiscal de otorgar o no la prórroga.
traslado_organo	id_organo	Numeric(19,0)	No	Identificador de los órganos de traslado.
	Órgano	Varchar(255)	No	Tipo de órgano al que se traslada.

Tabla 6. Nivel de granularidad para el subsistema VF.

2.4.7 Identificación de las dimensiones

Las dimensiones representan cada uno de los ejes en un espacio multidimensional. Suministran el contexto en el que se obtienen las medidas de un hecho. Las dimensiones se utilizan para seleccionar y agrupar los datos en un nivel de detalle deseado. Los componentes de una dimensión se denominan niveles y se organizan en jerarquías.

Constituyen las perspectivas de análisis de la información y presentan entre sus características principales la definición de jerarquías entre sus atributos, cuyo objetivo es modelar explícitamente la forma en que se puede consolidar el proceso de análisis de la información. En el desarrollo de la solución se identificaron las siguientes dimensiones:

Para el subsistema CLEP

- **Dimensión tema:** tratamiento penitenciario, laboral, libertades anticipadas, materia penal, situación legal o ejecutoria.
- **Dimensión término:** describe si una acción determinada fue resuelta en término o fuera de término.
- **Dimensión traslado órgano:** contiene los órganos a los que puede ser trasladada una denuncia como son el órgano de intrusión, Fiscalía militar y PNR.
- **Dimensión tipo de escrito:** define el tipo de escrito que llega a la fiscalía como queja, petición o denuncia.
- **Dimensión instancia:** definiría en que instancia de las fiscalías existentes de nuestro país se están obteniendo los datos, las cuales serían Fiscalía General de la Republica, Provincial y Municipal.
- **Dimensión persona:** contiene la raza y el sexo de la persona que ha presentado un escrito.
- **Dimensión decisión depurar:** contiene la decisión que se toma en la depuración que puede ser archivar, trasladar o tramitar.
- **Dimensión prórroga:** define las prórrogas solicitadas al fiscal.
- **Dimensión fallo del tribunal:** describe si las impugnaciones fueron declaradas con lugar o sin lugar.
- **Dimensión procedencia:** contiene la procedencia de los dictámenes.
- **Dimensión tipo de resolución:** contiene el tipo de resolución ya sea un auto una sentencia una resolución etc.
- **Dimensión tiempo:** esta dimensión situaría cada acción realizada en la fiscalía en el tiempo, se define esta dimensión con varias categorías o niveles para su mejor organización, estos son año, trimestre, mes y día.
- **Dimensión decisión dictamen:** describe si la decisión es favorable o desfavorable.
- **Dimensión tipo de dictamen:** contiene la descripción del tipo de dictamen que puede ser libertad condicional, libertad condicional extraordinaria, salida excepcional del territorio nacional, entre otras.
- **Dimensión tipo subsidiaria:** contiene el tipo de subsidiaria que puede ser TCCI, TSCI, LL, entre

otras.

- **Dimensión provincia:** contiene el nombre de las provincias del país y el identificador de cada uno de ellos.
- **Dimensión decisión tramitar:** describe si la decisión de la tramitación.

Para el subsistema PDC

- **Dimensión instancia:** definiría en que instancia de las fiscalías existentes de nuestro país se están obteniendo los datos, las cuales serían Fiscalía General de la Republica, Provincial y Municipal.
- **Dimensión tiempo:** esta dimensión situaría cada acción realizada en la fiscalía en el tiempo, se define esta dimensión con varias categorías o niveles para su mejor organización, estos son año, trimestre, mes y día.
- **Dimensión decisión adoptada:** contiene la decisión adoptada por la fiscalía.
- **Dimensión traslado penal:** contiene los traslados penales realizados.
- **Dimensión revisión:** contiene las revisiones realizadas.
- **Dimensión disposición fiscal:** esta dimensión contiene si se realiza un escrito promoción o un dictamen denegatorio.
- **Dimensión persona:** contiene el sexo de la persona involucrada.
- **Dimensión fallo del tribunal:** describe si las acciones realizadas fueron declaradas con lugar o sin lugar.
- **Dimensión tipo de pruebas:** contiene información sobre los tipos de pruebas realizadas en la FGR.

Para el subsistema VF

- **Dimensión tipo de responsable:** contiene el tipo de responsable de las medidas disciplinarias.
- **Dimensión instancia:** definiría en que instancia de las fiscalías existentes de nuestro país se están obteniendo los datos, las cuales serían Fiscalía General de la Republica, Provincial y Municipal.

- **Dimensión traslado órgano:** contiene los órganos a los que puede ser trasladada una denuncia como son el órgano de intrusión, Fiscalía militar y PNR.
- **Dimensión tiempo:** esta dimensión contiene el tiempo perteneciente a un hecho ocurrido, se define esta dimensión con varias categorías o niveles para su mejor organización, estos son año, mes y día.
- **Dimensión categoría:** contiene la categoría de los hechos.
- **Dimensión fallo del tribunal:** describe si las acciones realizadas fueron declaradas con lugar o sin lugar.
- **Dimensión acciones iniciadas:** contiene información sobre las acciones de control que son iniciadas.
- **Dimensión acciones concluidas:** contiene información sobre las acciones de control que son concluidas.
- **Dimensión violaciones:** contiene la información de las diferentes violaciones ocurridas.
- **Dimensión medidas disciplinarias:** en esta dimensión se almacenan los datos necesarios de las medidas disciplinarias aplicadas.
- **Dimensión termino:** contiene si una acción determinada fue resuelta en término o fuera de término.
- **Dimensión aprobación prórroga:** almacena la decisión del fiscal de otorgar o no la prórroga.

2.4.8 Identificación de las tablas de hechos

El hecho es el objeto a analizar, posee atributos nombrados de hechos o de síntesis y son generalmente de tipo cuantitativo. Sus valores (medidas) se obtienen generalmente por la aplicación de una función estadística que resume un conjunto de valores en un único valor. Para el desarrollo de la solución se identificaron los siguientes hechos:

Para el subsistema CLEP

- **Hecho registro de quejas:** en esta tabla se recogerán todas las medidas relacionadas con los

registros de escritos.

- **Hecho rollo de quejas:** en esta tabla se guardan todas las medidas relacionadas con la tramitación de las quejas.
- **Hecho dictámenes:** en esta tabla se recogerán todas las medidas relacionadas con los dictámenes.

Para el subsistema VF

- **Hecho acciones:** en esta tabla se guardan las medidas relacionadas con las acciones de control.
- **Hecho hechos:** en esta tabla se guardan todas las medidas relacionadas con los hechos delictivos.

Para el subsistema PDC

- **Hecho solicitudes recibidas:** en esta tabla se almacenan las medidas relacionadas con las solicitudes recibidas en la fiscalía.
- **Hecho rollo de tramitación:** en esta tabla se almacenan las medidas relacionadas con las revisiones penales en tramitación.
- **Hecho rollo FGR:** es esta tabla se almacenan algunos procesos elevados a la FGR.

2.4.9 Identificación de atributos de dimensiones y tablas de hechos

En las Tabla 7, 8 y 9 se muestran todas las tablas de dimensiones con sus respectivos atributos.

Nombre de la dimensión	Atributos
dim_tiempo	id_tiempo, dia, mes, anno, trimestre
dim_instancia	id_fiscalia, descripcion
dim_traslado_organo	id_organo, órgano
dim_tipo_escrito	id_tipo_escrito, descripcion
dim_tema	id_tema, descripcion
dim_persona	id_persona, sexo, raza

dim_decision_depurar	id_decision_dictamen, descripcion
dim_termino	id_termino, cant_dias
dim_prorroga	id_documento, descripcion
dim_fallo_tribunal	id_fallo, descripcion
dim_procedencia	id_procedencia, descripcion
dim_tipo_subsidiaria	id_tipo_subsidiaria, descripcion
dim_tipo_dictamen	id_tipo_dictamen, descripcion
dim_decision_dictamen	id_decision_dictamen, descripcion
dim_tipo_resolucion	id_tipo_resolucion, descripcion
dim_decision_tramitar	id_decision_tramitar, discriminante

Tabla 7. Atributos de las tablas dimensiones del subsistema CLEP.

Nombre de la dimensión	Atributos
dim_tiempo	id_tiempo, dia, mes, anno, trimestre
dim_instancia	id_fiscalia, descripcion
dim_persona	id_persona, sexo
dim_fallo_tribunal	id_fallo, descripcion
dim_decision_adoptada	id_decision_adop, descripcion
dim_traslado_penal	id_traslado, descripcion
dim_revision	id_revision, descripcion

dim_disposicion_fiscal	id_disposicion, decision_fiscal
dim_tipo_prueba	id_tipo_prueba, descripcion

Tabla 8. Atributos de las tablas dimensiones del subsistema PDC.

Nombre de la dimensión	Atributos
dim_tiempo	id_tiempo, dia, mes, anno, trimestre
dim_instancia	id_fiscalia, descripcion
dim_traslado_organo	id_organo, órgano
dim_tipo_responsable	id_tipo_responsable, descripcion
dim_aprobacion_prorroga	id_documento, decisión
dim_fallo_tribunal	id_fallo, descripcion
dim_accion_iniciada	id_documento, descripcion
dim_acciones_concluidas	id_documento, descripcion
dim_violaciones	id_violaciones, causas_condiciones
dim_medidas_disciplinarias	id_med_solicitada, discriminante
dim_categoria	id_categoria, descripcion

Tabla 9. Atributos de las tablas dimensiones del subsistema VF

En las Tablas 10, 11 y 12 se muestran las tablas de hechos con sus atributos:

Nombre de los hechos	Atributos

hecho_registro_escrito	id_fiscalia, id_registro_escrito, id_tipo_escrito, id_organo, id_tema, id_persona, id_decision_depurar
hecho_rollo_de_quejas	id_termino, id_rollo_queja, id_fiscalia, id_prorroga, id_impugnacion, id_decision_tramitar, id_fallo
hecho_dictámenes	id_termino, id_dictámenes, id_procedencia, id_fiscalia, id_tipo_subsidiaria, id_tiempo, id_tipo_dictamen, id_decision_dictamen, id_persona, id_tipo_resolucion, id_fallo

Tabla 10. Atributos de las tablas hechos del subsistema CLEP.

Nombre de los hechos	Atributos
hecho_acciones	id_documento, id_tiempo, id_documento, id_fiscalia, id_documento
hecho_hechos	id_fiscalia, id_violaciones, id_med_solicitadas, id_categoria, id_hechos, id_tipo_responsable, id_tiempo

Tabla 11. Atributos de las tablas hechos del subsistema VF.

Nombre de los hechos	Atributos
hecho_solicitudes_recibidas	id_fiscalia, id_solicitud id_tiempo, id_decision_adop, id_traslado_penal, id_persona
hecho_rollos_tramitacion	id_rollo_tramitacion, id_revision, id_fiscalia, id_tiempo, id_fallo, id_disposicion
hecho_rollo_FGR	id_fiscalia, id_rollo_fgr id_tiempo, id_tipo_prueba

Tabla 12. Atributos de las tablas hechos del subsistema PDC.

2.4.10 Patrones de diseño de Bases de Datos para los MD

Los patrones constituyen una solución estándar para un problema común.

Los patrones de diseño de una Base de Datos permiten al usuario crear una BD más fortalecida ya que constituyen una guía que especifica cómo debe ser la misma. En la actualidad las Bases de Datos suelen ser muy grandes y a veces el trabajo con los patrones de diseño hacen que el diseño sea más fácil y que se asegure un resultado satisfactorio.

2.4.10.1 Patrón de llaves subrogadas

Este patrón es muy utilizado pues se decide generar una llave primaria única para cada entidad en vez de usar un atributo identificador en el contexto dado. (18) Esto permite que las tablas sean más fáciles de consultar a partir del identificador, pues todos tienen el mismo tipo en cada una de las tablas. En los Mercados de Datos realizados, las tablas de dimensiones tienen una llave primaria única al igual que las tablas de hechos, con la diferencia de que las tablas de hechos toman como llave foránea a las llaves primarias de las dimensiones asociadas a él.

2.4.10.2 Patrón de constelación de hechos

El modelo constelación de hechos es más complejo que las otras arquitecturas debido a que contiene múltiples tablas de hechos. En este las tablas de dimensiones pueden estar compartidas con más de una tabla de los hechos. Este modelo es bastante flexible, sin embargo, el problema es que cuando el número de las tablas vinculadas aumenta, la arquitectura puede ser muy compleja y difícil para poderlo mantener. (19) Este patrón está presente en los tres Mercados de Datos realizados y se muestra en los diseños lógicos que se realizaron a cada uno de ellos, como se muestra en la Figura 7 y en los Anexos 2 y 3.

2.4.11 Estandarización del código

La estandarización es el proceso mediante el cual se realiza una actividad de manera previamente establecida. El término estandarización proviene del término estándar, aquel que refiere a un modo o método establecido, aceptado y normalmente seguido para realizar determinado tipo de actividades o funciones. Un estándar es un parámetro más o menos esperable para ciertas circunstancias o espacios y es aquello que debe ser seguido en caso de recurrir a algunos tipos de acción. El término de estandarización tiene como connotación principal la idea de seguir entonces el proceso estándar a través del cual se tiene que actuar o proceder.

Seguidamente se explican los estándares de código a seguir para la realización de los Mercados de Datos del subsistema PDC, CLEP y VF del SIGEF:

- Los nombres de las entidades y atributos serán con minúscula, separando por “_” cuando exista composición de palabras.
- Las palabras que tenga ñ esta serán tratadas como “nn”.
- Las tablas hechos comenzarán con la palabra “hecho”.
- Las tablas dimensiones comenzarán con “dim”.
- Los identificadores de las tablas comenzarán su nombre con “id”.
- Las palabras que lleven tilde se pondrán sin estas en las tablas.

2.4.12 Tipo de modelo lógico del Mercado de Datos

La variante de modelamiento que se propone utilizar en el desarrollo de los Mercados de Datos para el subsistema PDC, CLEP y VF es **Esquema Constelación de hechos**, debido a las necesidades de los usuarios. Se selecciona este esquema porque existirán tablas de dimensiones que se relacionarán con más de una tabla de hechos.

2.4.13 Arquitectura de los Mercados de Datos

La arquitectura definida para el desarrollo de la solución consta de dos subsistemas: almacenamiento e integración, ubicados en dos niveles y la fuente de datos como se muestra en la Figura 6.



Figura 6. Arquitectura del sistema.

En el primer nivel se encuentra el subsistema de integración, el cual se abastece de las diferentes fuentes de datos y se encarga de llevar a cabo los procesos que extraen, integran y transforman la información para su almacenamiento. Los usuarios que acceden a este subsistema son los encargados de la administración de dichos procesos.

Por su parte, el subsistema de almacenamiento recibe la información manipulada durante la extracción, transformación y carga, finalmente se almacena en una Base de Datos soportada por el SGBD PostgreSQL y administrada por los usuarios autorizados mediante la herramienta PgAdmin III.

2.4.14 Modelo lógico de los datos

Una vez definidas dentro del negocio las dimensiones, hechos y medidas se procede a la estructuración de

los modelos dimensionales correspondientes a cada Mercado de Datos.

En la Figura 7 se muestra el modelo de datos lógico diseñado para el subsistema PDC, donde se evidencia el uso de la topología constelación de hechos, atendiendo a que existen varias tablas de hechos que comparten algunas de sus dimensiones.

En la Anexo 2 se muestra el modelo de datos lógico diseñado para el subsistema CLEP, donde se evidencia el uso de la topología constelación de hechos, atendiendo a que existen varias tablas de hechos que comparten algunas de sus dimensiones.

En la Anexo 3 se muestra el modelo de datos lógico diseñado para el subsistema VF, donde se evidencia el uso de la topología constelación de hechos, atendiendo a que existen varias tablas de hechos que comparten algunas de sus dimensiones.

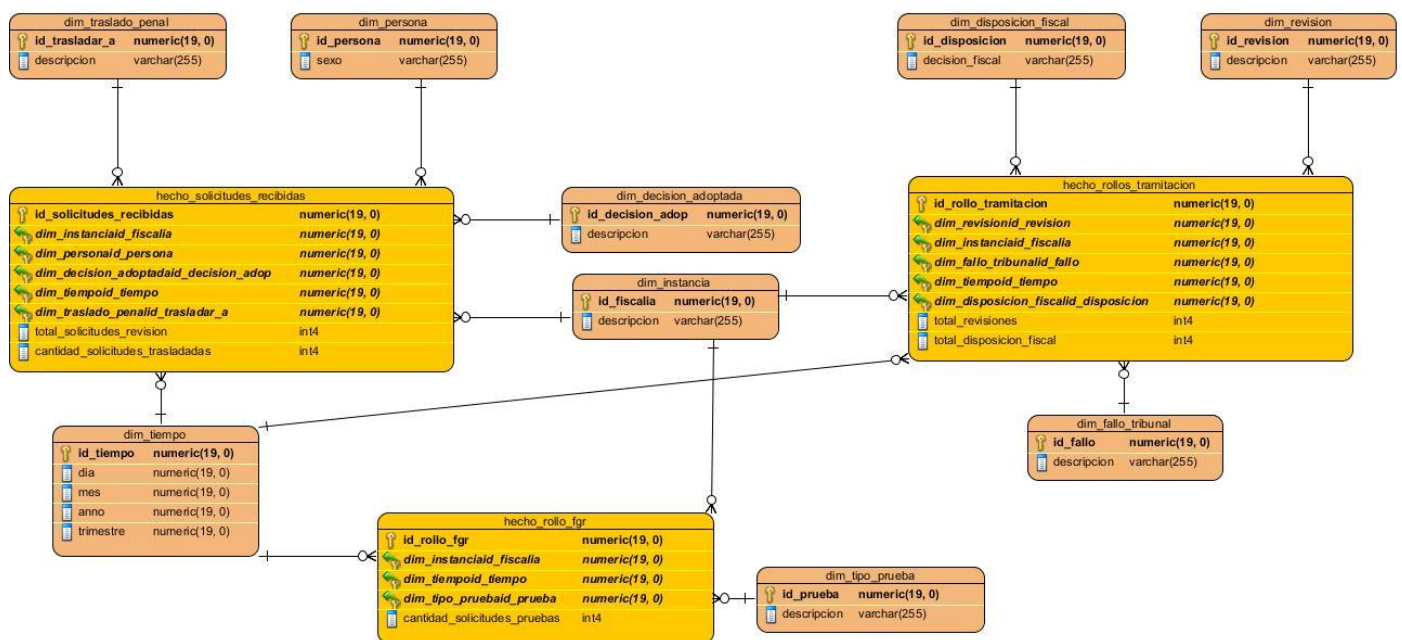


Figura 7. Modelo lógico de datos para el subsistema PDC.

2.4.15 Modelo físico de los datos

Una vez realizado el modelo lógico de los Mercados de Datos se realiza el modelo físico para cada uno de ellos como se muestra en la Figura 8, 9 y 10.

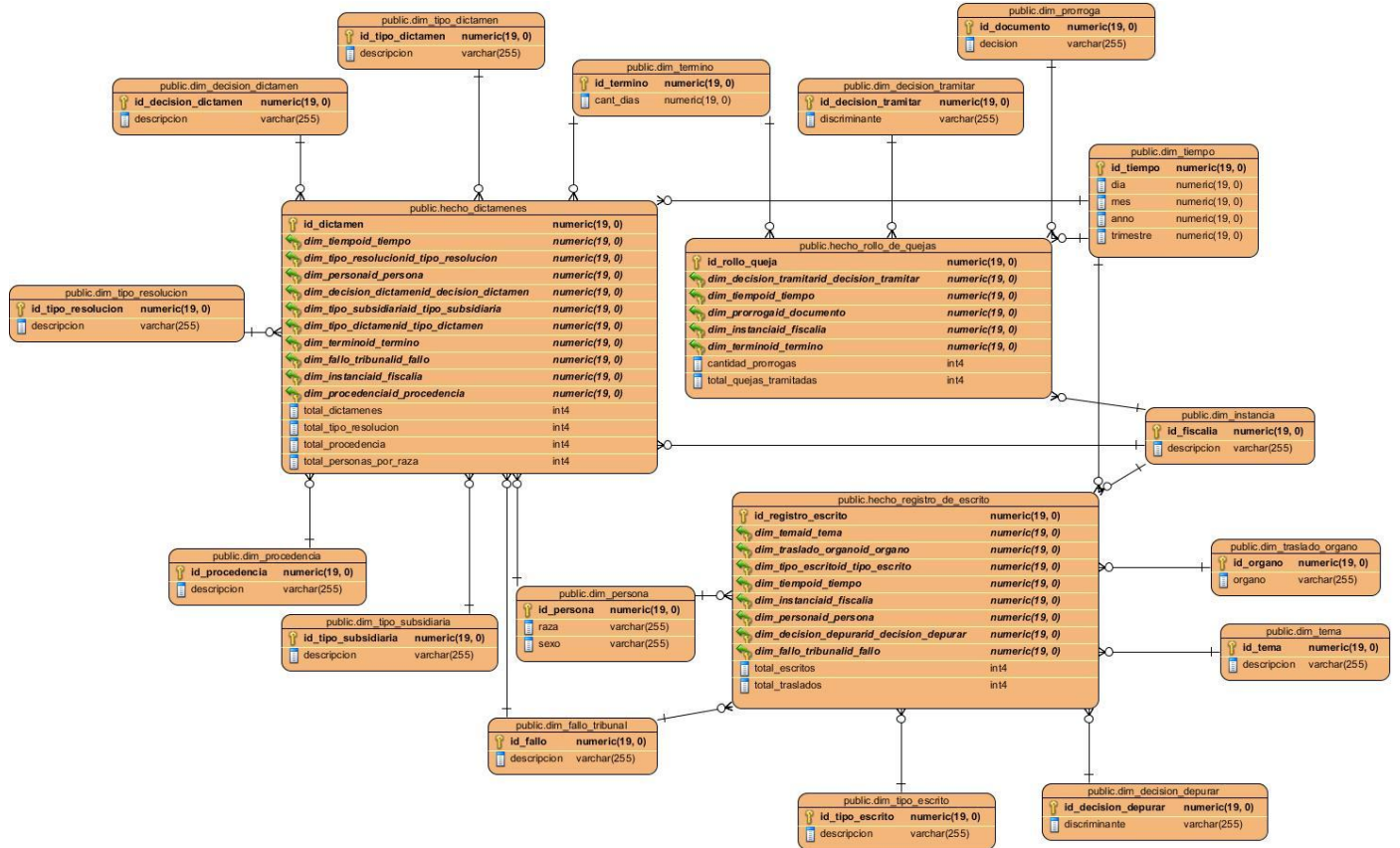


Figura 8. Modelo físico de datos para el subsistema CLEP.

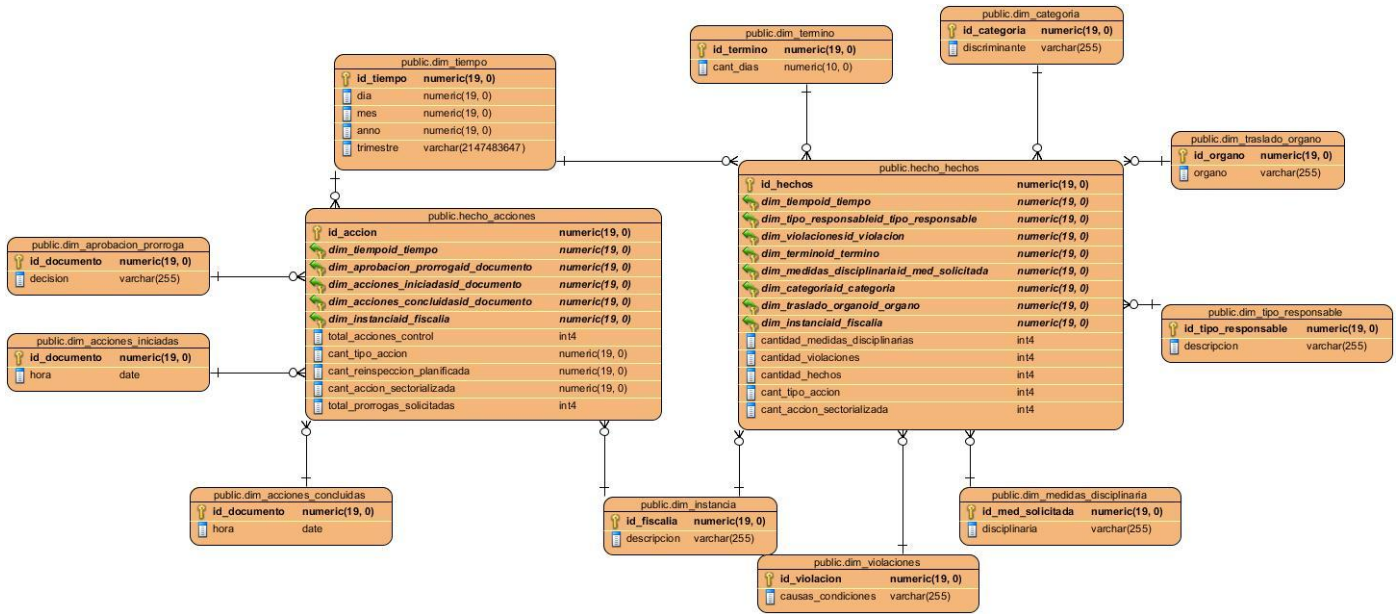


Figura 9. . Modelo físico de datos para el subsistema VF.

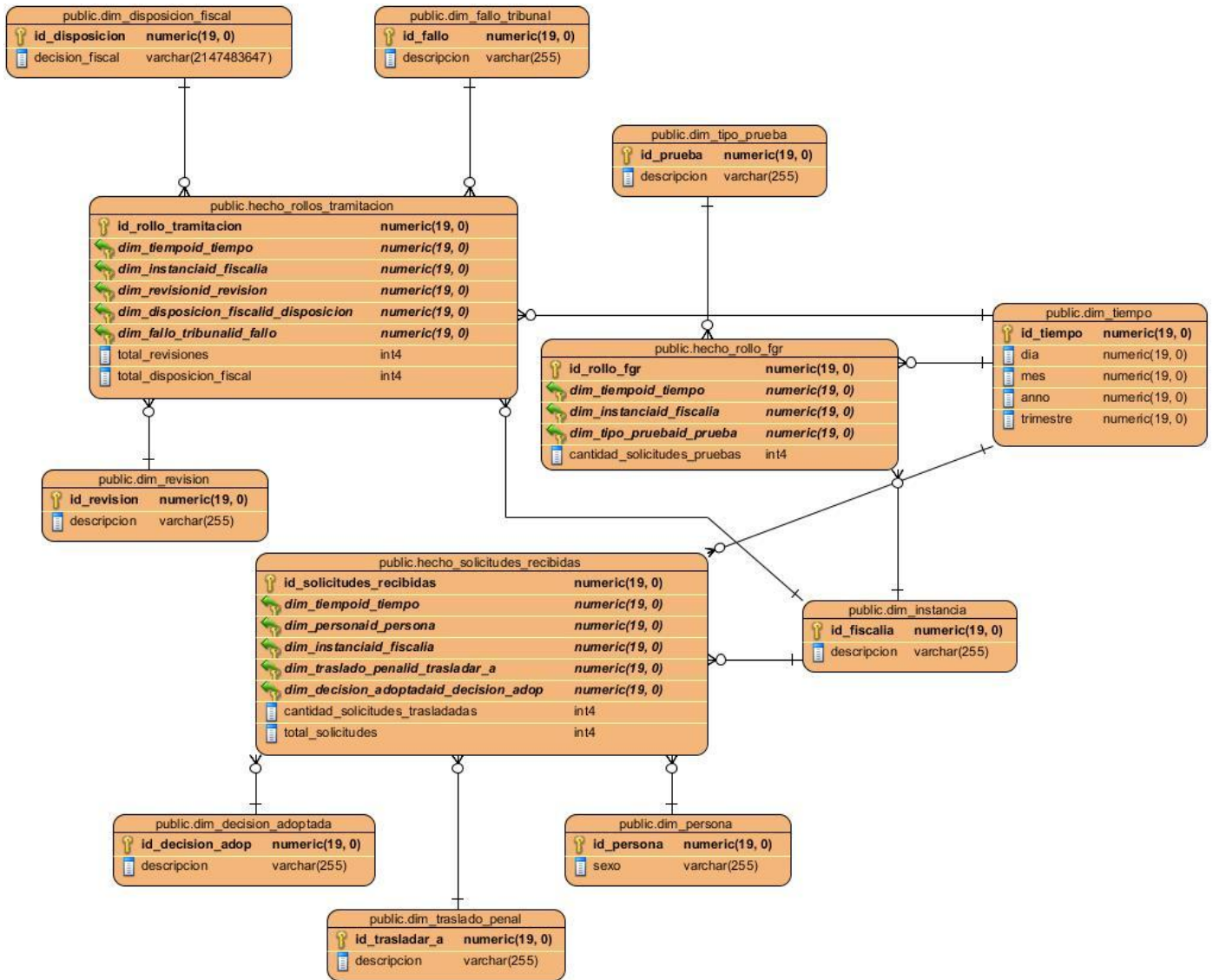


Figura 10. Modelo físico de datos para el subsistema PDC.

2.4.16 Prueba del modelo

Se le realizan pruebas al modelo haciendo preguntas para ver si puede responder a estas, un ejemplo de esta sería:

¿Cuántos escritos recibidos en la fiscalía son quejas?

¿Cuántas denuncias han sido trasladadas a la PNR?

¿Cuántas quejas fueron archivadas?

Como se puede observar en la Figura 11 la tabla hecho_registro_escrito tiene relación con las dimensiones: decision_depurar, traslado_organismo, tipo_escrito; además contiene las medidas total_quejas_archivadas, total_traslados_pnr y total_quejas_recibidas, por lo que el modelo está apto para responder a estas preguntas.

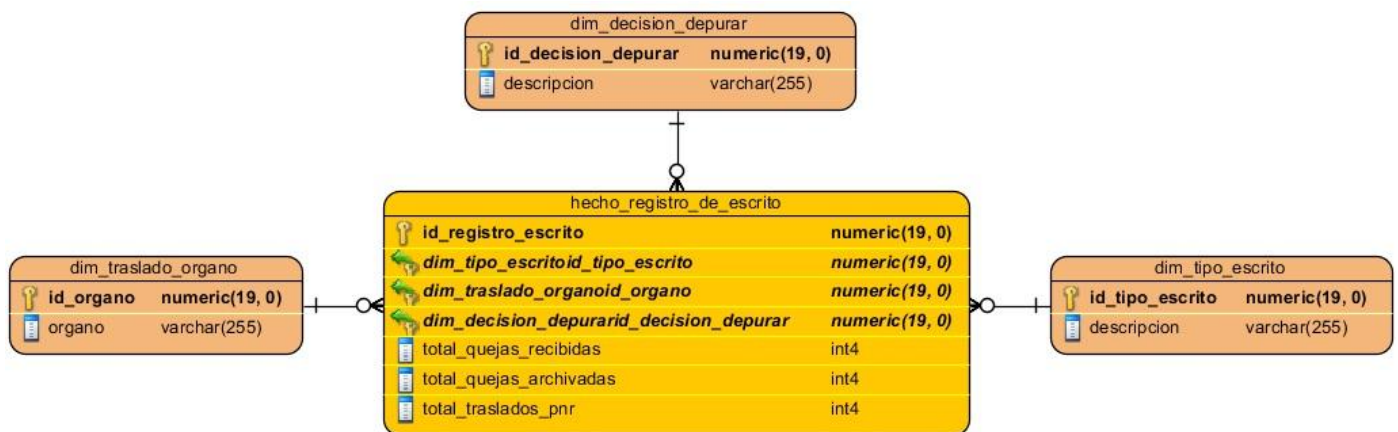


Figura 11. Fragmento del modelo conceptual.

2.4.17 Diseño de los sistemas de extracción, transformación y carga de los datos.

Antes de comenzar el diseño de las transformaciones de los datos fue necesario realizar una configuración de la conexión a la Base de Datos fuente y otra a la Base de Datos Almacén, estas conexiones se pueden observar en el Anexo 4 y luego se pasa a hacer las transformaciones.

Las transformaciones están compuestas por pasos enlazados entre sí a través de los saltos. Los pasos constituyen el elemento más pequeño dentro de las transformaciones, y a través de los saltos fluye la información entre los diferentes pasos. Algunos de los pasos más comunes que se realizaron son: entrada de tabla, e insertar/actualizar. En los Anexos del 5 al 7 se puede observar los componentes. En la Figura 12 se muestra el diseño de la transformación para la dimensión tiempo que es una de las dimensiones que se encuentra repetida en los tres Mercados de Datos y en la Figura 13 se muestra la transformación del hecho rollo de quejas que pertenece al Mercado de Datos del subsistema CLEP en la Figura 14 se muestra la

transformación del hecho rollo de FGR perteneciente al Mercado de Datos del subsistema PDC y en la Figura 15 se muestra la transformación realizada al hecho acciones del Mercado de Datos del subsistema VF. A partir del Anexo 14 hasta el Anexo 44 se muestran las transformaciones realizadas a las dimensiones y hechos y desde el Anexo 45 hasta el Anexo 51 se pueden consultar las consultas realizadas para cada uno de los hechos realizados para los Mercados de Datos.



Figura 12. Transformación para la dimensión "dim_tiempo."



Figura 13. Transformación para el hecho "hecho_rollo_queja".



Figura 14. Transformación para el hecho "hecho_rollo_fgr".



Figura 15. Transformación para el hecho "hecho_acciones".

Basándose en las necesidades de información de la fiscalía para los subsistemas CLEP, VF y PDC, se explora la fuente de datos a disposición, y se extrae la información que se considere relevante. Para la realización de este proceso se crearon un conjunto de consultas SQL que dan cumplimiento a este objetivo. Las mismas se incorporan al componente de entrada de Kettle.

Por último se cargan los datos hasta los Mercados de Datos y se le realiza actualizaciones o mantenimientos periódicos, esto se realiza a partir de las transformaciones previamente diseñadas. Para dar cumplimiento a la misma se diseñó un trabajo para cada tabla de hecho, donde primero se cargarán los datos de las dimensiones asociadas a este y luego se cargan los de la tabla de hecho. También fue necesario crear un

trabajo general denominado “general” por cada Mercado de Datos donde se realiza una carga inicial de todos los datos. Esta carga inicial se refiere precisamente a la primera carga de datos que se realiza a los Mercados de Datos, por lo general esta tarea consume bastante tiempo, ya que se deben de insertar todos los registros que han sido generados.

2.5 Implementación de los trabajos

En el contexto de integración de datos, el término trabajo o job en inglés, es un conjunto de tareas que se realizan con el objetivo de ejecutar una acción determinada. La implementación de un trabajo define una secuencia lógica para la ejecución de las transformaciones, mediante el uso de pasos definidos, los cuales son diferentes a los disponibles en las transformaciones. Es posible ejecutar una o varias transformaciones de las que se hayan diseñado y organizar una secuencia de ejecución para ellas. Los trabajos se encuentran en un nivel superior a las transformaciones.

En esta investigación se realizaron tres trabajos generales, uno para cada Mercado de Datos, estos trabajos se conectan primeramente a la Base de Datos y verifican la conexión, si están conectados comienzan con la carga, de lo contrario terminan su ejecución. La carga comienza con las transformaciones de las dimensiones compartidas por los hechos y luego carga cada uno de los trabajos de los hechos, que estos contienen las dimensiones propias y la transformación del hecho en sí. En los Anexos 8, 9 y 10 se muestran imágenes de estos trabajos generales.

2.6 Diseño de las dimensiones y cubo de información

Luego de haber terminado el diseño de las transformaciones, trabajos y haber cargado los datos a sus Mercados de Datos correspondientes, se establece una conexión a los Mercados de Datos, y posteriormente se realiza el diseño de las dimensiones y los cubos de información en la herramienta Schema-Worbench.

Primeramente se diseñan los cubos multidimensionales y dentro de estos las medidas y dimensiones correspondientes a cada uno definiendo para esta última las jerarquías, niveles y propiedades. Las dimensiones no se crean en un cubo específicamente, estas se crean fuera, para que otros cubos puedan utilizarlos si llega a ser necesario y dentro de los cubos se les hace referencia a las que utiliza cada cual.

A continuación se muestra la creación multidimensional en el Schema-Worbench.

1. Se crean las dimensiones.
2. Se crean los cubos.

Una vez diseñadas todas las dimensiones y cubos pasamos a publicar los Mercados de Datos para poder ver los resultados. En los Anexos 11, 12 y 13 se muestran los esquemas realizados en el Schema-Workbench para cada uno de los Mercados de Datos.

2.7 Visualización de los resultados

Luego de publicar los Mercados de Datos con la ayuda de la herramienta biserver-ce que viene incluida en el paquete del Pentaho, podemos visualizar los resultados obtenidos. En las Figuras 13 y en los Anexos 53 y 54 se muestran las representaciones de un cubo de datos publicado por cada uno de los Mercados de Datos realizados para los subsistemas CLEP, PDC y VF:

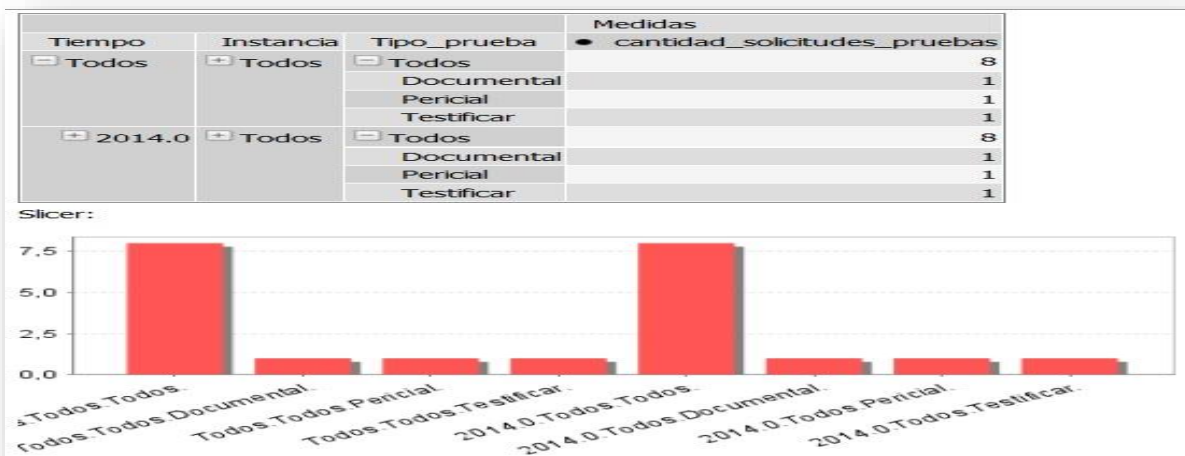


Figura 16. Representación en el navegador del cubo Rollo_FGR.

2.8 Conclusiones del capítulo

En el presente capítulo se abordó sobre las etapas de diseño e implementación realizadas a los Mercados de Datos para los subsistemas CLEP, PDC y VF y una vez terminadas estas etapas se arrojaron las siguientes conclusiones:

- Se realizó un estudio de las necesidades de información del cliente que permitió la identificación de las dimensiones y hechos con sus principales criterios de medidas correspondientes a cada uno de los Mercados de Datos.
- Mediante el diseño de los modelos de datos fueron identificadas diecisiete tablas dimensionales y tres tablas de hechos para el subsistema CLEP, diez tablas dimensionales y dos tablas de hechos para el subsistema VF y para el subsistema PDC fueron identificadas nueve tablas de dimensiones y tres tablas de hechos, que garantizan el correcto funcionamiento del sistema.
- Se realizaron en total ocho transformaciones para la carga de los hechos, treinta para las dimensiones y tres para los trabajos generales, que posibilitaron la carga de los datos de la fuente.

3 Capítulo 3. Pruebas

3.1 Introducción

Para garantizar el cumplimiento de las exigencias del cliente y la calidad del producto se realizan las pruebas de validación. Este capítulo contiene las diferentes técnicas de pruebas que se emplearon, así como el análisis de los resultados obtenidos. Se realizaron las pruebas de volumen y carga, además de pruebas de carga y estrés, las cuales validarán la utilización de los Mercados de Datos.

3.2 Pruebas

Las pruebas constituyen un elemento crítico para la calidad del software. De acuerdo a la IEEE una prueba se define como: Actividad en la cual un sistema o componente es ejecutado bajo condiciones específicas, se observan o almacenan los resultados y se realiza una evaluación de algún aspecto del sistema o componente. (20) Para comprobar el funcionamiento de los Mercados de Datos se realizaron dos tipos de pruebas, las pruebas de Volumen y Carga y las de Carga y Estrés.

3.3 Pruebas de carga y estrés

Para la realización de las pruebas de carga y estrés se utiliza la herramienta JMeter en su versión 2.4 por la facilidad de su uso y las funcionalidades que brinda. Estas pruebas son aplicadas con el objetivo de obtener datos sobre la carga del sistema que ayuden a revisar el dimensionamiento del sistema, generando carga en el mismo mediante la simulación de concurrencia que se acerque con fiabilidad a lo esperado en la explotación real. A continuación se muestra la cantidad de usuarios que tendrán los Mercados de Datos en los distintos niveles de la FGR.

Instancia	Cantidad de usuarios	
Fiscalía General de la República	25 (real)	50(prueba)
Fiscalía Provincial	12(real)	25(prueba)

El rango de tiempo aceptable para la ejecución de las consultas, queda definido de la siguiente forma:

Tipo de respuesta	Tiempo
Admisible	Menor o igual que 5 segundos
No Admisible	Mayor que 5 segundos

Estas pruebas consisten en realizar una consulta y ver el tiempo de ejecución de consultas para una cantidad de usuarios conectados concurrentemente.

Seguidamente en la Figura 17 y 18 se muestran los resultados obtenidos para el Mercado de Datos correspondiente al subsistema PDC.

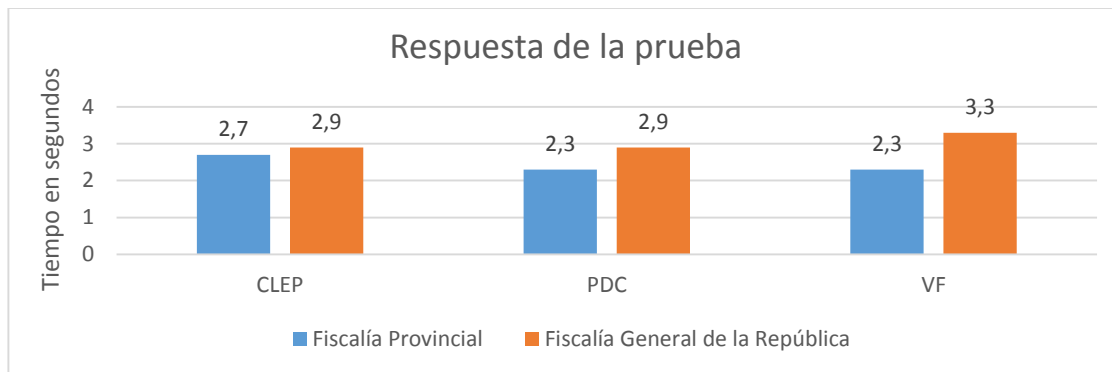
Etiqueta	# Muestras	Media	Mín	Máx	Desv. Estándar	% Error	Rendimiento
Petición JDBC2	500	320	2	3042	832,69	0,00%	148,1/sec
Total	500	320	2	3042	832,69	0,00%	148,1/sec

Figura 17. Resultado de la prueba de carga simulando cantidad de usuarios de la Fiscalía General.

Etiqueta	# Muestras	Media	Mín	Máx	Desv. Estándar	% Error	Rendimiento
Petición JDBC2	250	260	1	2497	730,37	0,00%	92,9/sec
Total	250	260	1	2497	730,37	0,00%	92,9/sec

Figura 18. Resultado de la prueba de carga simulando cantidad de usuarios de la Fiscalía Provincial.

Una vez realizadas estas pruebas de carga a las Bases de Datos de los mercados se llega a la conclusión de que las mismas poseen el rendimiento adecuado para explotar su potencial en entornos de desarrollos reales. Las operaciones de consultas de datos se realizan en tiempos aceptables para su correcto funcionamiento, además con el régimen de trabajo propuesto las Bases de Datos de los mercados se exponen a un nivel muy bajo de ocurrencia de errores. La gráfica que se muestra a continuación se realiza partiendo del tiempo máximo de ejecución de una consulta en cada uno de los Mercados de Datos según los resultados obtenidos por la herramienta JMeter.



Gráfica 1. Tiempos de respuesta máximos en segundos para 25 y 50 usuarios.

3.4 Pruebas de Volumen y Carga

Las pruebas de volumen y carga consisten en agregar gran cantidad de datos a los Mercados de Datos, para comprobar si es posible alcanzar un límite que haga fallar el software. También identifica la carga máxima continua o volumen que el elemento de prueba es capaz de manipular en un período de tiempo determinado. Las pruebas de volumen brinda la posibilidad de confirmar que la aplicación esté funcionando correctamente con el tamaño máximo esperado de Base de Datos.

Para realizar las pruebas de carga se utilizó la herramienta Datanamic Data Generator para PostgreSQL, la misma permitió llenar los Mercados de Datos con una determinada cantidad de datos validando si el servidor cuenta con los suficientes recursos de almacenamiento. Se configuró esta generación con valores arbitrarios, pero coincidentes en cuanto a su tipo y volumen.

Para probar los Mercados la cantidad de valores que se insertó en cada uno de ellos fue de 10 000 tuplas por tabla. Para definir esta cantidad se tuvo en cuenta que anualmente se espera que ocurra un aproximado de 20000 procesos entre todas las fiscalías del país, los Mercados de Datos responden a las necesidades de información del subsistema PDC, CLEP Y VF, de los módulos Verificación Fiscal, Quejas, Peticiones y Denuncias, Dictámenes y Revisiones de causas Penales, se guardan en el los datos de los procesos que están relacionados a las necesidades de información de los mismos. Se estiman que en los procesos que abarcan los Mercados de Datos se produzcan un total de 25 000 tuplas anuales, que en un periodo de 5 años (tiempo de vida del Mercado de Datos) alcanzaría las 400 000 tuplas. Transcurrido los 5 años se procede a realizar una salva de los datos, se limpian los Mercados y se procede a comenzar a cargarlos

nuevamente. En las Figuras 19 y en los Anexos 55 y 56 se muestra el resultado de la prueba de volumen realizada en la herramienta por cada Mercado de Datos.

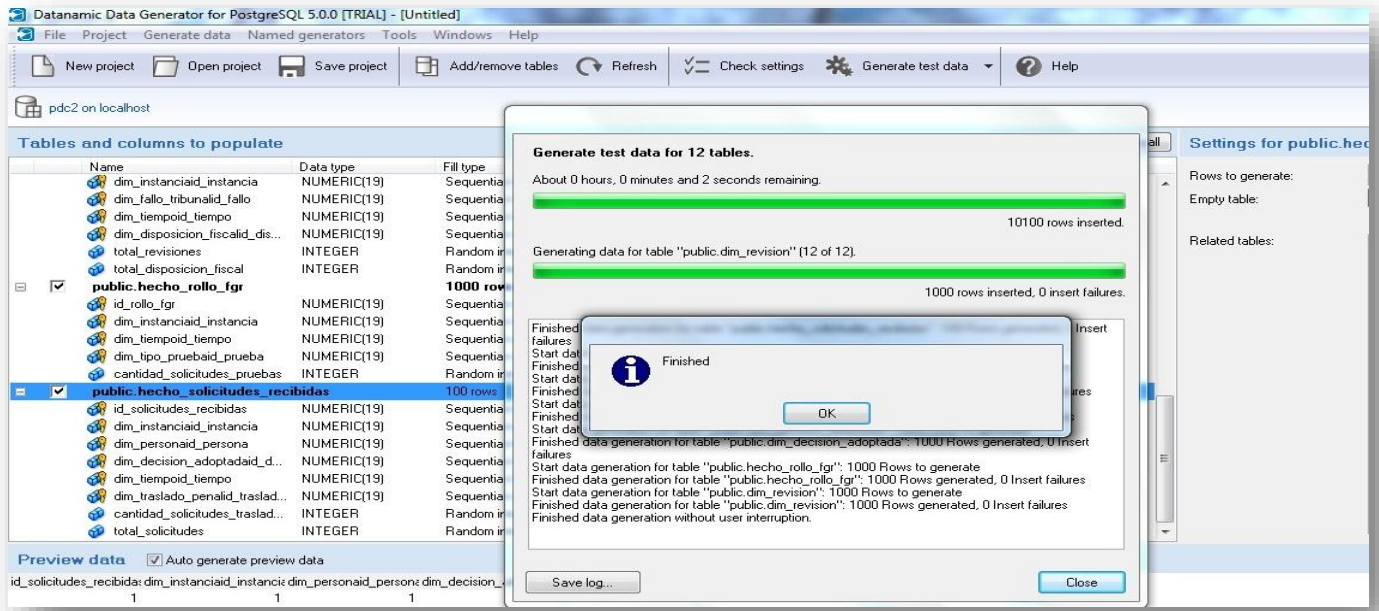


Figura 19. Resultado de la prueba de Carga en Datanamic Data Generator para el mercado del subsistema PDC.

Cuando se introducen los datos no se presentaron problemas de límite de capacidad, ni se detectaron desbordamientos de columnas, atributos, tipos de datos, ni peticiones excesivas de memoria.

Las llaves autogeneradas no se salieron del rango especificado, ni se detectaron problemas con los tipos de datos definidos en el paso de diseño. Lo anteriormente planteado garantiza que el SGBD utilizado y el diseño implementado portan completamente el almacenamiento de los niveles de información requeridos para la puesta en producción de los Mercados de Datos para PDC, CLEP Y VF.

Con la Base de Datos cargada con un volumen alto de datos se realizó la generación de carga con la herramienta JMeter, quedando definido la cantidad de usuarios y el tiempo de respuesta de la siguiente forma:

Instancia	Cantidad de usuarios
Fiscalía General de la República	25
Fiscalía Provincial	12

Tipo de respuesta	Tiempo
Admisible	Menor o igual que 5 segundos
No Admisible	Mayor que 5 segundos

Seguidamente en la Figura 20 y 21 se muestran los resultados obtenidos para el Mercado de Datos correspondiente al subsistema PDC.

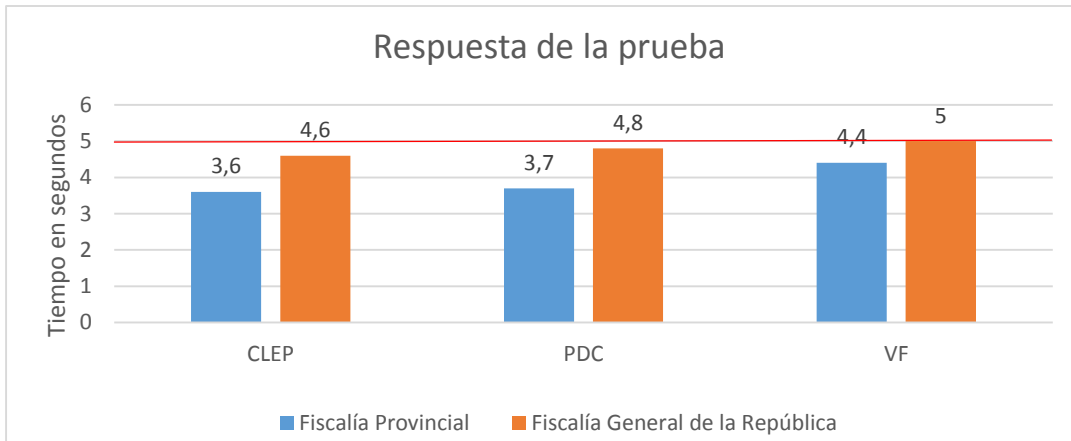
Etiqueta	# Muestras	Media	Mín	Máx	Desv. Estándar	% Error	Rendimiento
Petición JDBC2	500	720	16	4315	844,71	0,00%	58,0/sec
Total	500	720	16	4315	844,71	0,00%	58,0/sec

Figura 20. Resultado de la prueba de carga simulando cantidad de usuarios de la Fiscalía General.

Etiqueta	# Muestras	Media	Mín	Máx	Desv. Estándar	% Error	Rendimiento
Petición JDBC2	250	405	21	2986	749,79	0,00%	45,2/sec
Total	250	405	21	2986	749,79	0,00%	45,2/sec

Figura 21. Resultado de la prueba de carga simulando cantidad de usuarios de la Fiscalía Provincial.

La gráfica que se muestra a continuación se realiza partiendo del tiempo máximo de ejecución de una consulta en cada uno de los Mercados de Datos según los resultados obtenidos por la herramienta JMeter luego de haber cargado la Base de Datos.



Gráfica 2. Valores de respuesta máximos luego de la prueba de volumen y carga.

3.5 Conclusiones del capítulo

- En este capítulo se validó la solución propuesta mediante las técnicas de pruebas que se aplicaron a los Mercados de Datos, estas pruebas fueron las pruebas de volumen y carga y las pruebas de carga y estrés.
- Se obtuvieron los resultados correctos esperados, o sea, los Mercados de Datos están aptos para recibir el volumen de datos que se espera.
- Los Mercados de Datos son capaces de mantener el máximo de conexiones estimadas de manera concurrente. También se midió el tiempo de ejecución de las consultas y cumplen con los umbrales de aceptación definidos por los clientes.

Conclusiones Generales

- El análisis de los fundamentos teóricos relacionados con los Mercados de Datos permitió llegar a la conclusión que estos puede ayudar en una primera etapa a resolver la necesidad de información que tienen las empresas para mejorar el proceso de toma de decisiones.
- El estudio de las metodologías y herramientas realizado, garantizó que la metodología seleccionada guiara el proceso de desarrollo de los Subsistemas de almacenamiento e integración a través de cada etapa del ciclo de vida y la selección de las herramientas y tecnologías para la construcción de la solución propuesta cumplen con la política de migración a software libre, dando soporte a las necesidades del equipo de desarrollo.
- Se analizaron las diferentes necesidades de información que posee el cliente, para identificar las dimensiones y hechos, señalando como dimensiones importantes la dimensión tiempo y la dimensión instancia, además se identifican las uniones establecidas entre estos elementos, que forman los diseños de los Mercados de Datos.
- La implementación de los ocho esquemas, las ocho transformaciones para la carga de los hechos, treinta y siete para las dimensiones y once para los trabajos, permitió la integración de los datos históricos y su almacenamiento.
- Las pruebas efectuadas durante las distintas etapas de desarrollo permitieron comprobar la funcionalidad de los Mercados de Datos realizados. Los resultados obtenidos durante las últimas pruebas realizadas fueron satisfactorios, validando el cumplimiento de los objetivos propuestos.
- Se obtuvieron los Mercados de Datos para los departamentos Verificación Fiscal, Quejas, Peticiones y Denuncias, Dictámenes y Revisiones de Causas Penales, permitiendo agilizar el análisis de los datos generados por el SIGEF, favoreciendo el proceso de toma de decisiones en la institución.

Recomendaciones

Con el objetivo de mejorar la solución planteada se proponen las siguientes recomendaciones:

- Continuar el desarrollo de la solución implementando Mercados de Datos para los restantes módulos del SIGEF.
- Integrar los Mercados de Datos realizados en un Almacén para todo el sistema.
- Utilizar la metodología y herramientas seleccionadas en la investigación.

Referencias bibliográficas

1. **Bitam.** Business Intelligence. *Bitam*. [En línea] 2002. [Citado el: 23 de 11 de 2013.] <http://www.bitam.com/spanish/AcercadeBI.htm>.
2. *Data Warehouse para una empresa de transporte turístico*, . **Emma R Rizo Rizo, Yoel L Hernández Lantigua.** Cuba : Revista Retos Turísticos, 2010.
3. **Inmon, William H.** *Building the Data Warehouse*. Indianapolis : Wiley Publishing, 2005. ISBN-13: 978-0-7645-9944-6.
4. **Ralph Kimball, Margy Ross.** *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*. New York : John Wiley & Sons, Inc, 2002. ISBN 0-471-20024-7.
5. **Informática Hoy.** [En línea] 2012. www.informatica-hoy.com.ar.
6. **otros, Ralph Kimball y.** *The Data Warehouse Lifecycle Toolkit*. [aut. libro] **Ralph Kimball.** Wiley : ISBN-10: 0470149779, 2008.
7. **Mavilio, Alfredo Sifontes.** *Diseño del Datamart del subsistema de Conducidos*. 2009.
8. **ITATÍ, PAOLA.** [En línea] 2010. <https://docs.google.com/viewer?a=v&q=cache:dCGluOH-PIJ:200.45.54.90/depar/areas/informatica/SistemasOperativos/PaolaMonog.pdf>.
9. **Pressman, Roger S.** *Ingeniería del Software. Un enfoque práctico*. s.l. : Quinta Edición, 2005.
10. **Visual Paradigm.** *Boost Productivity with Innovative and Intuitive Technologies*. [En línea] 18 de 09 de 2013. [Citado el: 14 de 10 de 2013.] <http://www.visualparadigm.com/product/vpum/>.
11. **PostgreSQL . PostgreSQL .** [En línea] 2005. [Citado el: 13 de 12 de 2013.] <http://www.postgresql.org/about/news/309/>.
12. **Martínez, Rafael.** *PostgreSQL*. [En línea] 2 de 10 de 2013. [Citado el: 23 de 11 de 2013.] http://www.postgresql.org.es/sobre_postgresql.

13. summan. Soluciones y servicios de manejo documental e infraestructura informática. [En línea] 2008. <http://www.summan.com/pentaho/pentaho-bi-platform-server>.
14. Pentaho Corporation. Pentaho. Powerful Analytics Made Easy. [En línea] 2012. [Citado el: 16 de 12 de 2013.] <http://www.pentaho.com/explore/pentaho-data-integration/>.
15. Díaz, Josep Curto. *Introducción al Business Intelligence*. Barcelona : s.n., 2010. ISBN:978-84-9788-886-8.
16. ESPINOSA, ROBERTO. Pentaho Data Integration. [En línea] 2010. <http://churriwifi.wordpress.com/2010/06/01/comparativa-talend-vs-kettle-pdi>.
17. Apache Jmeter . *Apache* . [En línea] 2013. www.apache.org..
18. Blaha, Michael. *Patterns of Data Modeling*. 2010.
19. Esquema de constelación de hechos (fact constellation schema). [En línea] http://etl-tools.info/es/bi/almacenedatos_esquema-constelacion.htm.
20. IEEE-SA. *IEEE-SA*. [En línea] 2013. <http://standards.ieee.org/develop/project/29119-1.html>.
21. pgAdmin. PostgreSQL Tools. [En línea] 2013. [Citado el: 15 de 11 de 2013.] <http://www.pgadmin.org>.
22. Gravatar. Gravatar. [En línea] 2013. [Citado el: 02 de 12 de 2013.]
23. Martinez, Rafael. postgresql-es. [En línea] 2013. http://www.postgresql.org/es/sobre_postgresql.

Anexos

Anexo # 1

Para verificación fiscal

Promedio de violaciones detectadas por acción de control concluida.

Total de medidas disciplinarias solicitadas a directivos superiores por responsabilidad directa en verificaciones fiscales.

Total de medidas disciplinarias solicitadas a directivos por responsabilidad directa en verificaciones fiscales.

Total de medidas disciplinarias solicitadas a ejecutivos por responsabilidad directa en verificaciones fiscales.

Total de violaciones de la legalidad detectadas.

- cantidad en verificaciones fiscales.
- cantidad en investigaciones fiscales.

Total de violaciones resueltas al momento de las reinspecciones.

- cantidad asociadas a verificaciones fiscales.
- cantidad asociadas a investigaciones fiscales.

Cantidad de Planes de medidas recibidos por la Fiscalía.

- cantidad asociados a verificaciones fiscales.
- cantidad asociados a investigaciones fiscales.

Cantidad de planes de medidas devueltos por la Fiscalía.

- cantidad asociados a verificaciones fiscales.
- cantidad asociados a investigaciones fiscales.

Total de hechos delictivos detectados.

- cantidad asociados a verificaciones fiscales.
- cantidad asociados a investigaciones fiscales.

Total de hechos delictivos remitidos a la PNR.

- cantidad asociados a verificaciones fiscales.
- cantidad asociados a investigaciones fiscales.

Cantidad de acciones de control concluidas excediendo el término.

- cantidad de Verificaciones Fiscales.

- cantidad de Investigaciones.

Total de prórrogas solicitadas para la tramitación de las acciones de control.

- cantidad en Verificaciones Fiscales.
- cantidad en Investigaciones.

Total de prórrogas concedidas para la tramitación de las acciones de control.

- cantidad en Verificaciones Fiscales.
- cantidad en Investigaciones.

Total de reinspecciones realizadas.

- cantidad asociadas a verificaciones fiscales.
- cantidad asociadas a investigaciones fiscales.
- cantidad de reinspecciones sectorializadas.
- cantidad asociadas a verificaciones fiscales.
- cantidad asociadas a investigaciones fiscales.

Total de reinspecciones realizadas fuera de término.

- cantidad asociadas a verificaciones fiscales.
- cantidad asociadas a investigaciones fiscales.

De las reinspecciones realizadas fuera de término, cantidad de reinspecciones sectorializadas.

- cantidad asociadas a verificaciones fiscales.
- cantidad asociadas a investigaciones fiscales.
- Total de medidas disciplinarias solicitadas a trabajadores por responsabilidad directa en verificaciones fiscales.
- Total de medidas disciplinarias solicitadas a directivos superiores por responsabilidad colateral en verificaciones fiscales.
- Total de medidas disciplinarias solicitadas a directivos por responsabilidad colateral en verificaciones fiscales.
- Total de medidas disciplinarias solicitadas a ejecutivos por responsabilidad colateral en verificaciones fiscales.
- Total de medidas disciplinarias solicitadas a trabajadores por responsabilidad colateral en verificaciones fiscales.

- Total de medidas disciplinarias solicitadas a directivos superiores por responsabilidad directa en investigaciones fiscales.
- Total de medidas disciplinarias solicitadas a directivos por responsabilidad directa en investigaciones fiscales.
- Total de medidas disciplinarias solicitadas a ejecutivos por responsabilidad directa en investigaciones fiscales.
- Total de medidas disciplinarias solicitadas a trabajadores por responsabilidad directa en investigaciones fiscales.
- Total de medidas disciplinarias solicitadas a directivos superiores por responsabilidad colateral en investigaciones fiscales.
- Total de medidas disciplinarias solicitadas a directivos por responsabilidad colateral en investigaciones fiscales.
- Total de medidas disciplinarias solicitadas a ejecutivos por responsabilidad colateral en investigaciones fiscales.
- Total de medidas disciplinarias solicitadas a trabajadores por responsabilidad colateral en investigaciones fiscales.

Del total de violaciones resueltas al momento de las reinspecciones, cantidad de violaciones asociadas a sectorializadas.

- cantidad resueltas en término.
- cantidad resueltas fuera de término.

Del total de violaciones resueltas al momento de las reinspecciones, cantidad de violaciones resueltas fuera de término que corresponden a sectorializadas.

Total de violaciones pendientes de erradicar en reinspecciones.

- cantidad de violaciones que corresponden a sectorializadas.

De las violaciones pendientes de erradicar, las violaciones fuera de término.

- cantidad asociadas a verificaciones fiscales.
- cantidad asociadas a investigaciones fiscales.

De las violaciones pendientes de erradicar fuera de término, cantidad de violaciones que corresponden a sectorializadas.

Total de nuevas violaciones detectadas en reinspecciones.

- cantidad asociadas a verificaciones fiscales.
- cantidad asociadas a investigaciones fiscales.

Del total de nuevas violaciones detectadas en reinspecciones, cantidad de violaciones que corresponden a sectorializadas.

Total de Medidas disciplinarias solicitadas a cuadros responsables de las violaciones detectadas.

- cantidad asociadas a verificaciones fiscales.
- cantidad asociadas a investigaciones fiscales.

Total de medidas disciplinarias solicitadas a trabajadores responsables de las violaciones detectadas.

- cantidad asociadas a verificaciones fiscales.
- cantidad asociadas a investigaciones fiscales.

Total de medidas disciplinarias aplicadas a cuadros responsables de las violaciones detectadas.

- cantidad asociadas a verificaciones fiscales.
- cantidad asociadas a investigaciones fiscales.

Total de medidas disciplinarias aplicadas a trabajadores responsables de las violaciones detectadas.

- cantidad asociadas a verificaciones fiscales.
- cantidad asociadas a investigaciones fiscales.

Total de medidas disciplinarias aplicadas a cuadros responsables con las que el fiscal estuvo inconforme.

Total de medidas disciplinarias aplicadas a trabajadores responsables con las que el fiscal estuvo inconforme.

Total de medidas disciplinarias no aplicadas a cuadros responsables con lo que el fiscal estuvo inconforme.

Total de medidas disciplinarias no aplicadas a cuadros responsables con lo que el fiscal estuvo conforme.

Total de medidas disciplinarias no aplicadas a trabajadores responsables con lo que el fiscal estuvo inconforme.

Total de medidas disciplinarias no aplicadas a trabajadores responsables con lo que el fiscal estuvo conforme.

Del total de personas acusadas por los hechos delictivos que constituyen manifestaciones de corrupción, cantidad que son directivos superiores.

Del total de personas acusadas por los hechos delictivos que constituyen manifestaciones de corrupción, cantidad que son directivos.

Del total de personas acusadas por los hechos delictivos que constituyen manifestaciones de corrupción, cantidad que son ejecutivos.

Cantidad de procesos penales radicados.

Cantidad de procesos penales pendientes de radicar.

Total de personas acusadas por los hechos delictivos derivados de las acciones de control.

- cantidad asociadas a hechos delictivos remitidos a la PNR.
- cantidad asociadas a hechos delictivos remitidos al órgano de Instrucción.
- cantidad que son directivos superiores.
- cantidad que son directivos.
- cantidad que son ejecutivos.

Total de responsabilidades materiales solicitadas.

- cantidad asociadas a verificaciones fiscales.
- cantidad asociadas a investigaciones fiscales.

Total de responsabilidades materiales aplicadas.

- cantidad asociadas a verificaciones fiscales.
- cantidad asociadas a investigaciones fiscales.

Para el departamento revisiones penales:

Cantidad de solicitudes de pruebas documental.

Cantidad de solicitudes de pruebas testificar.

Cantidad de solicitudes de pruebas pericial.

Total de revisiones dictaminadas por departamento provincial en período.

Total de revisiones denegadas por el departamento provincial.

Total de solicitudes de revisión presentadas al tribunal que fueron declaradas con lugar.

Total de solicitudes de revisión presentadas al tribunal que fueron declaradas con lugar en parte.

Total de solicitudes de revisión presentadas al tribunal que fueron declaradas sin lugar.

Total de escrito de promoción.

Total de dictamen denegatorio.

Para el departamento Quejas, Peticiones y Denuncias:

Total general de quejas relacionadas con tratamiento penitenciario.

Total de quejas relacionadas con materia penal.

Total de quejas relacionadas con libertades anticipadas.

Total general de quejas laborales.

Total de quejas relacionadas Situación legal o ejecutoria.

Cantidad de impugnaciones declaradas con lugar.

Cantidad de impugnaciones declaradas sin lugar.

Cantidad de impugnaciones declaradas con lugar en parte.

Cantidad de prórrogas recibidas.

Para el departamento Dictámenes:

Total de autos del tribunal desfavorable por concepto de salida excepcional del territorio nacional presentados por la DEP.

Total de recursos de súplica interpuestos presentados por la DEP.

Total de casos presentados por la PNR para salida excepcional del territorio nacional.

Total de casos presentados por la PNR para salida excepcional del territorio nacional en TCCI.

Total de casos presentados por la PNR para salida excepcional del territorio nacional en TCSI.

Total de casos presentados por la PNR para salida excepcional del territorio nacional en LL.

Total de casos presentados por la PNR para salida excepcional del territorio nacional en remisión condicional de la sanción.

Total de casos presentados por la PNR para salida excepcional del territorio nacional en licencia extrapenal.

Total de casos presentados por la PNR para salida excepcional del territorio nacional en otras situaciones.

Total de dictámenes favorables del Fiscal por concepto de salida excepcional del territorio nacional presentado por la PNR.

Total de dictámenes desfavorables del Fiscal por concepto de salida excepcional del territorio nacional presentado por la PNR.

Total de autos del tribunal favorable por concepto de salida excepcional del territorio nacional presentados por la PNR.

Total de autos del tribunal desfavorable por concepto de salida excepcional del territorio nacional presentados por la PNR.

Total de recursos de súplica por concepto de salida excepcional del territorio nacional interpuestos por el Fiscal de los casos presentados por la PNR.

Total de solicitudes presentadas al Fiscal por concepto libertad condicional.

Total de dictámenes favorables del Fiscal por concepto de libertad condicional.

Total de dictámenes desfavorables del Fiscal por concepto de libertad condicional.

Total de autos del tribunal favorable por concepto de libertad condicional.

Total de autos del tribunal desfavorable por concepto de libertad condicional.

Total de recursos de súplica interpuestos por el Fiscal por concepto de libertad condicional.

Total de solicitudes presentadas al Fiscal por concepto libertad condicional extraordinaria.

Total de dictámenes favorables del Fiscal por concepto libertad condicional extraordinaria.

Total de dictámenes desfavorables del Fiscal por concepto libertad condicional extraordinaria.

Total de autos del tribunal favorables del Fiscal por concepto libertad condicional extraordinaria.

Total de autos del tribunal desfavorables del Fiscal por concepto libertad condicional extraordinaria.

Total de recursos de súplica interpuestos por el Fiscal por concepto de libertad condicional extraordinaria.

Total de sustitutivas dictaminadas favorables por el Fiscal presentadas por la DEP.

Total de substitutivas dictaminadas desfavorables por el Fiscal presentadas por la DEP.

Total de substitutivas dictaminadas favorables por el Fiscal presentadas por la PNR.

Total de substitutivas dictaminadas desfavorables por el Fiscal presentadas por la PNR.

Total de autos del tribunal favorable por concepto de salida excepcional del territorio nacional presentados por la DEP.

Total de substitutivas presentadas al tribunal.

Total de substitutivas aprobadas por el tribunal.

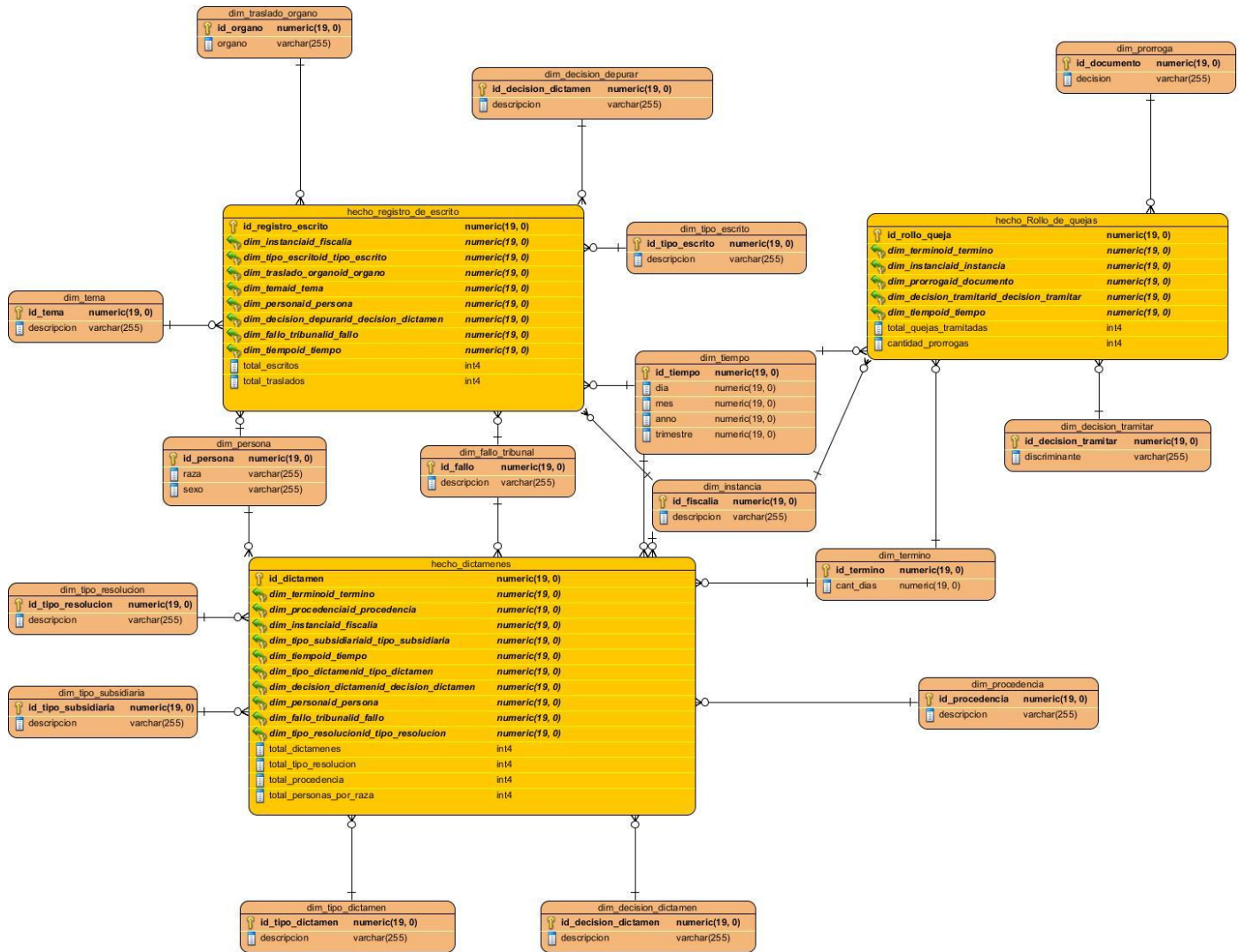
Total de substitutivas denegadas por el tribunal.

Total de libertades condicionales denegadas por el tribunal.

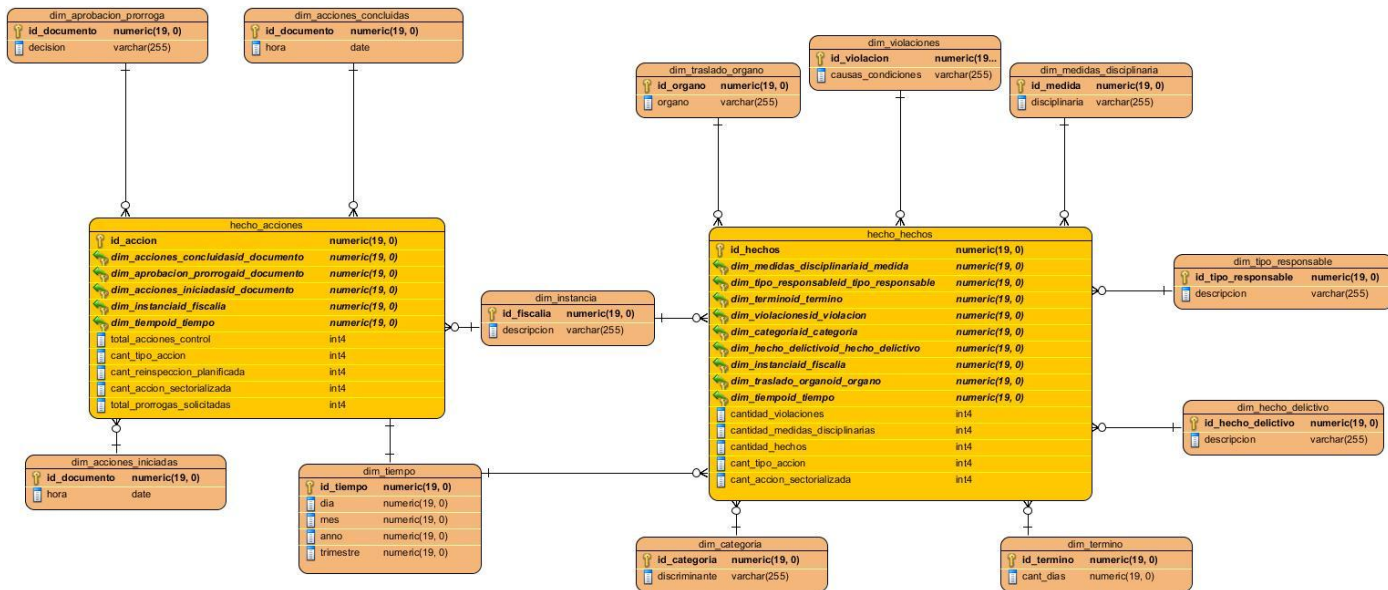
Total de recursos de súplica interpuestos por L/C.

Total de recursos de súplica declarados con lugar por L/C.

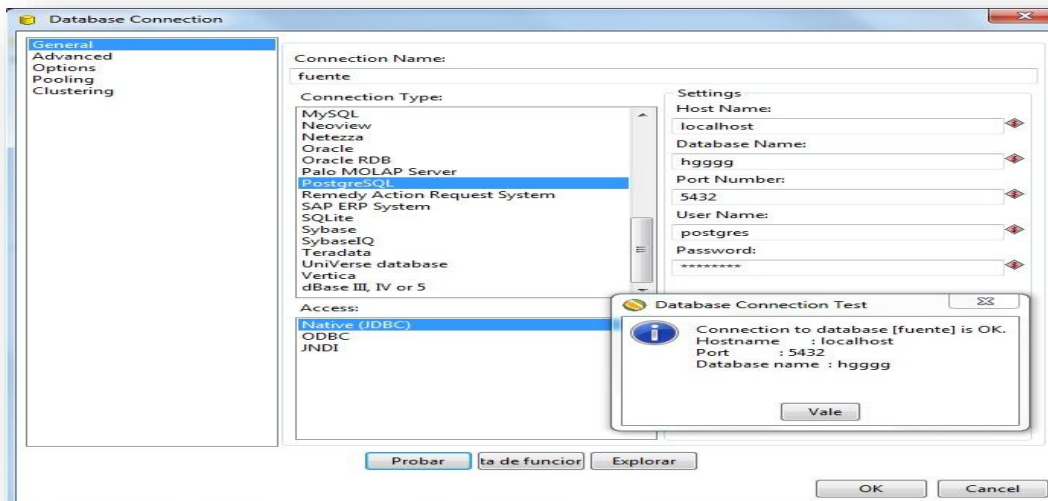
Total de recursos de súplica declarados sin lugar por L/C.



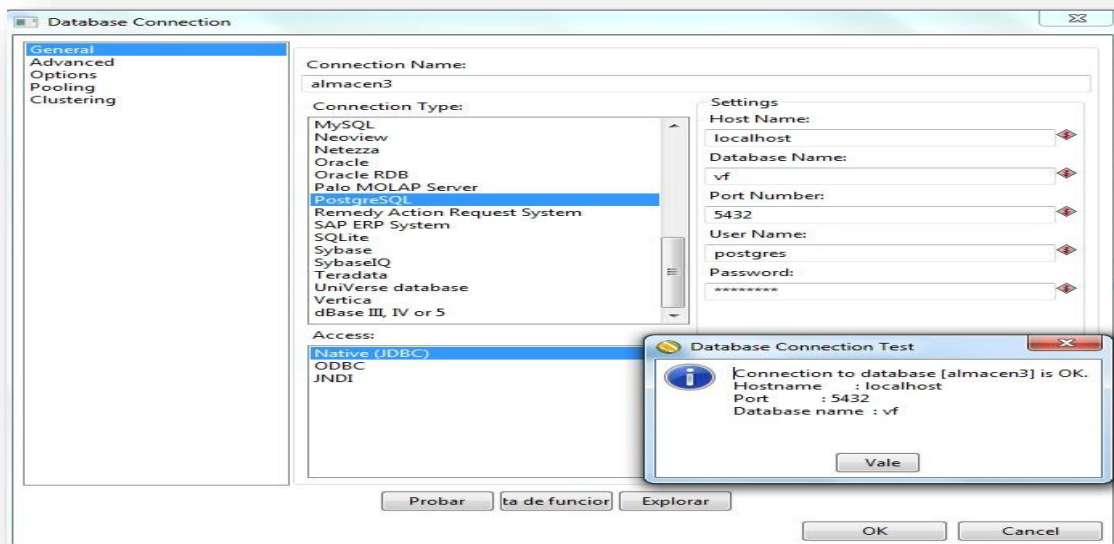
Anexo 2 Modelo de datos para el subsistema CLEP.



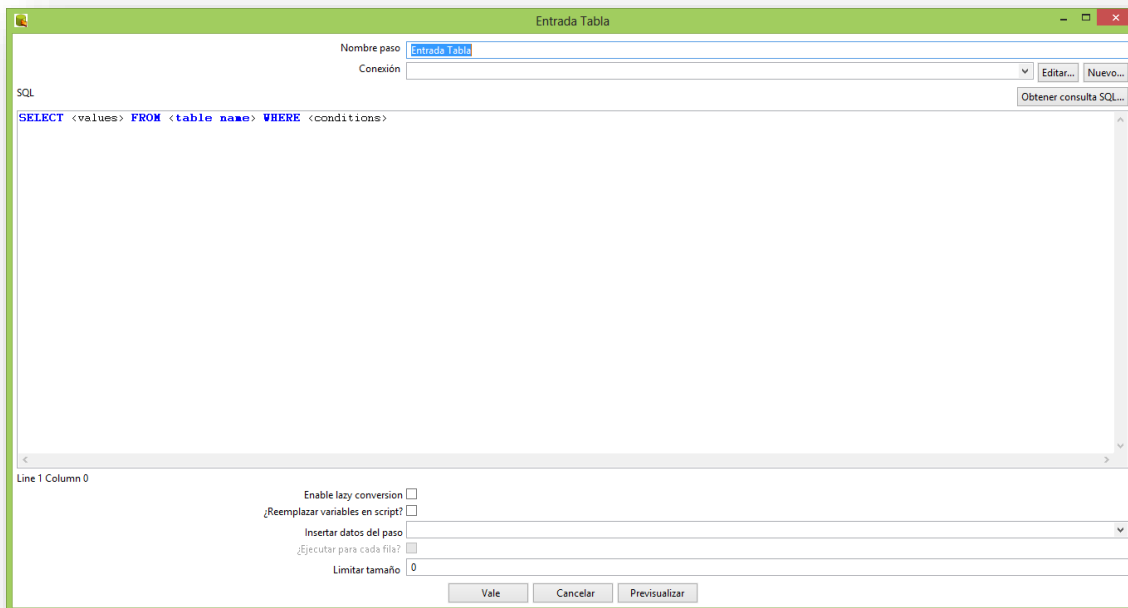
Anexo 3 Modelo de datos para el subsistema VF.



Anexo 4 Conexión con la Base de Datos fuente.



Anexo 4 Conexión con la Base de Datos destino.



Anexo 5: Componente para la entrada de datos

Nombre paso:

Conexión:

Esquema de búsqueda:

Tabla de búsqueda:

¿Habilitar cache?

Tamaño de cache en filas (0=todas):

Load all data from table

La clave(s) para realizar búsqueda de valor(es):

#	Campo de tabla	Comparador	Campo1	Campo2
1				

Valores a devolver de la tabla de búsqueda:

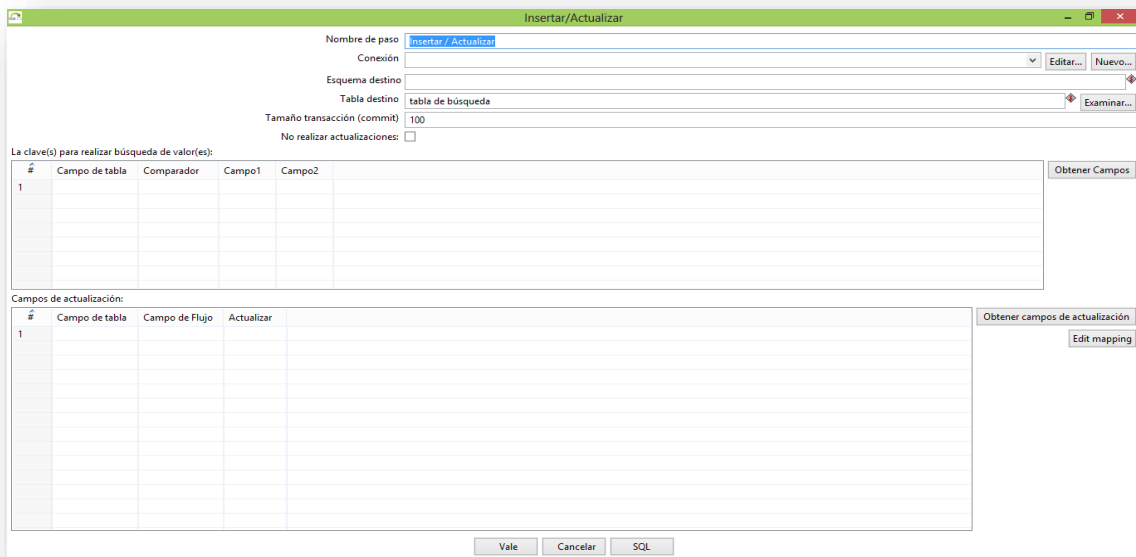
#	Campo	Nuevo nombre	Defecto	Tipo
1				

No procesar la fila si la búsqueda falla

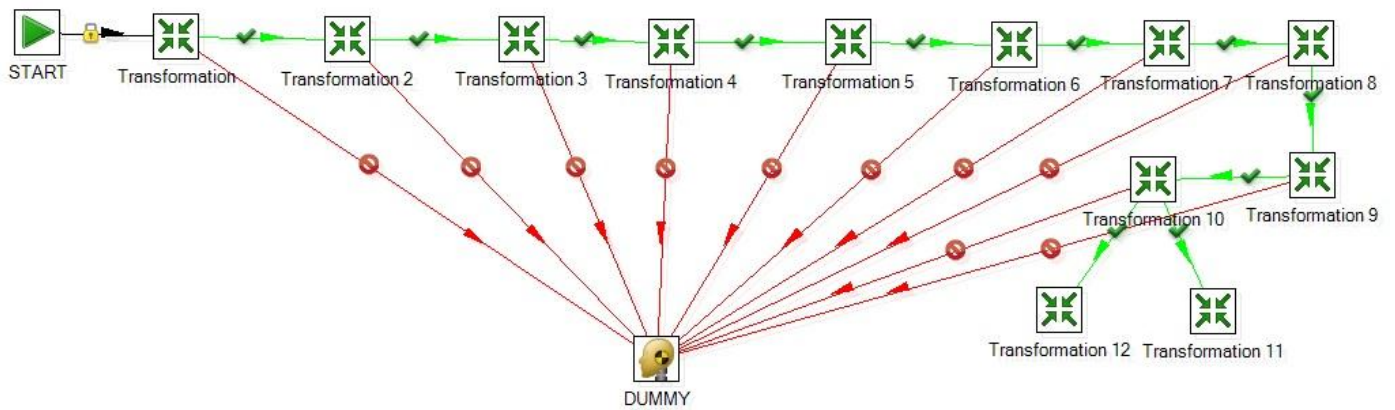
¿Producir error si se obtienen múltiples resultados?

Ordenar por:

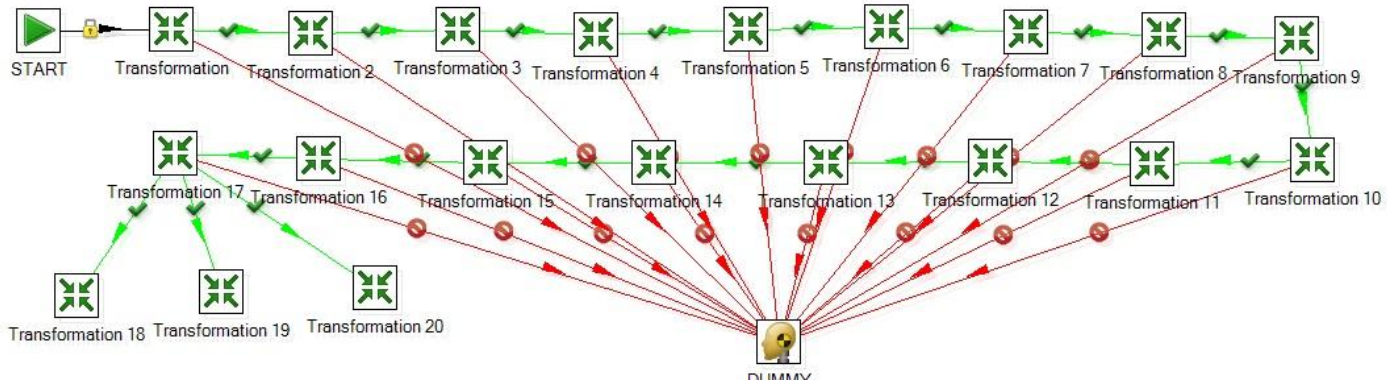
Anexo 6. Componente para realizar búsqueda en Base de Datos



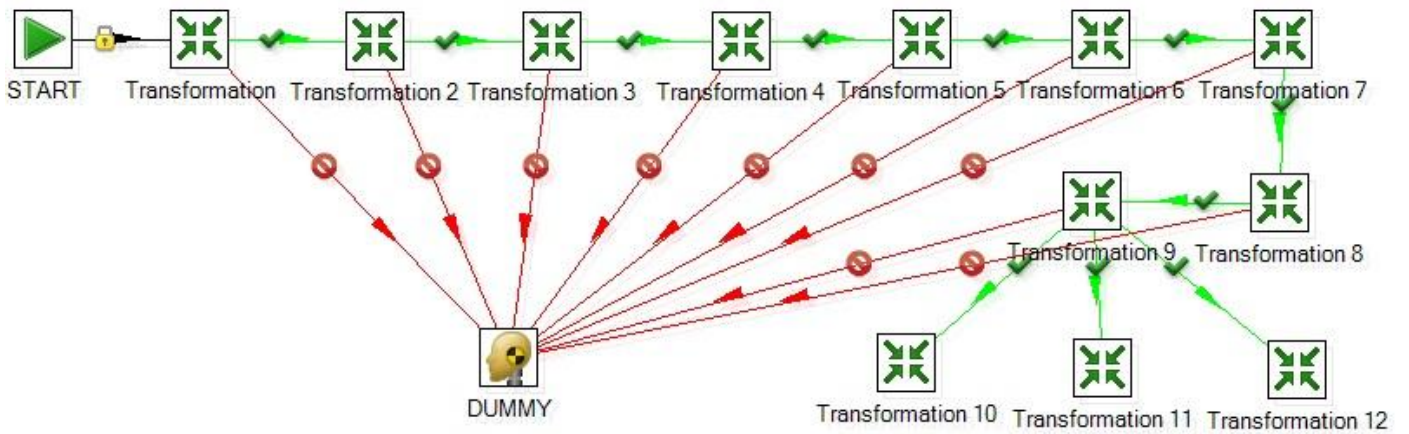
Anexo 7. Componente para insertar o actualizar los valores



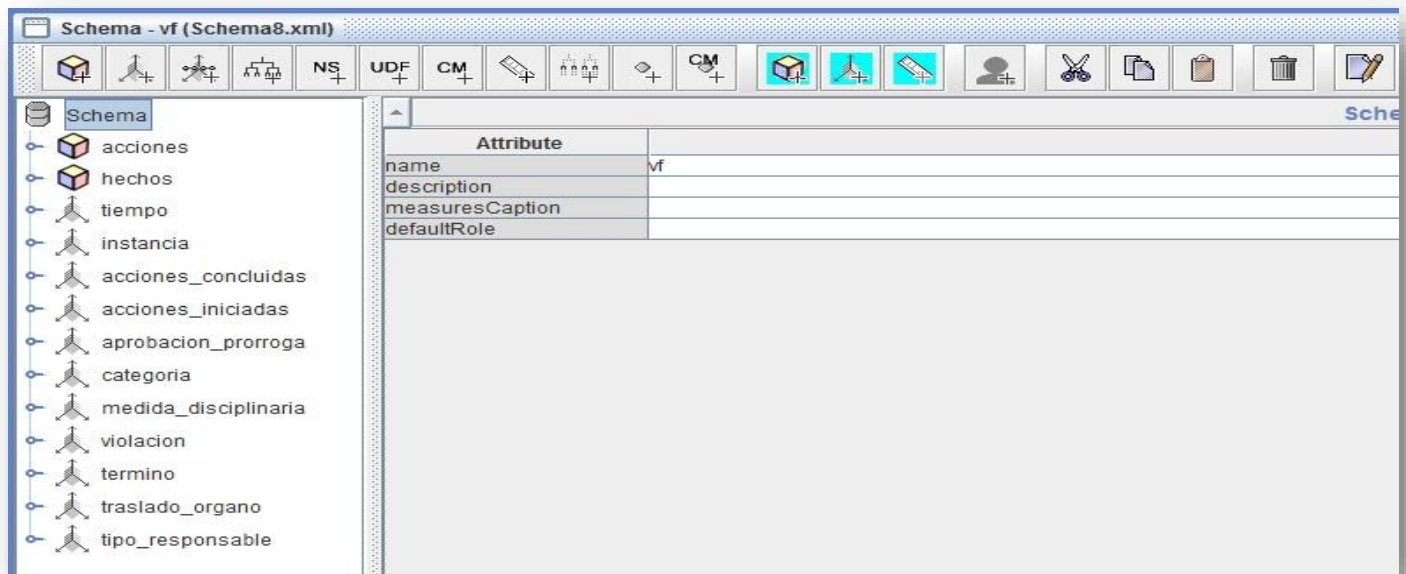
Anexo 8. Trabajo general para el Mercado de Datos del subsistema VF.



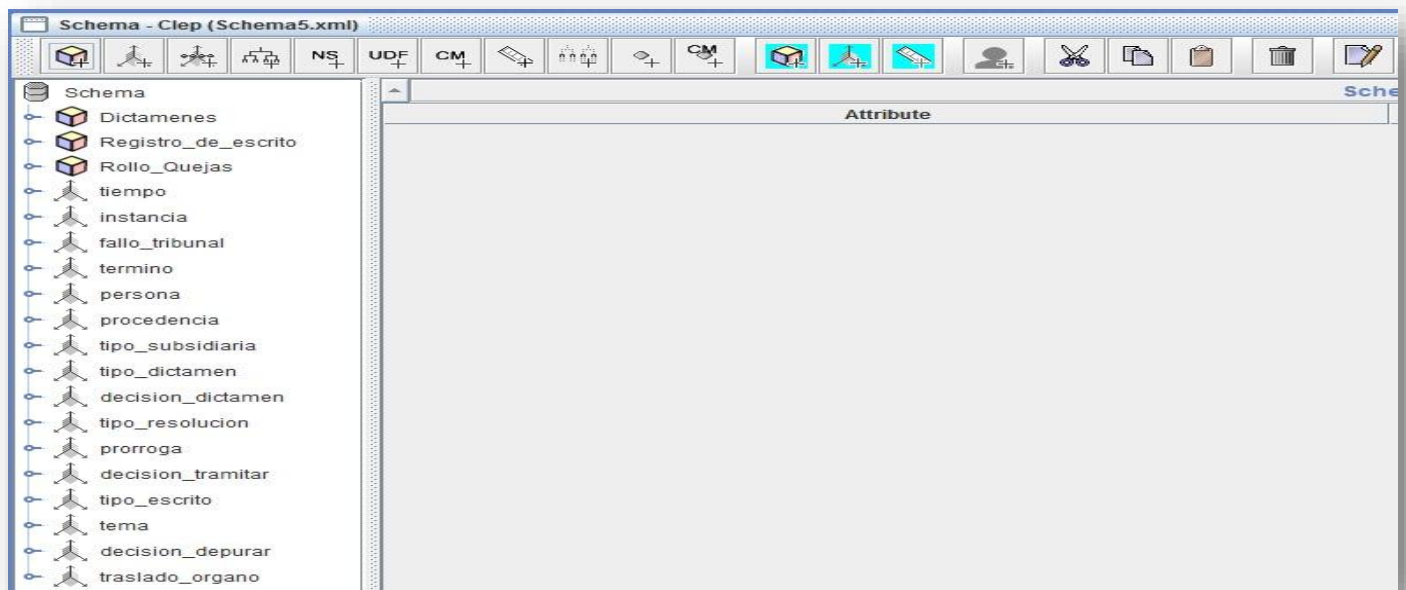
Anexo 9. Trabajo general para el Mercado de Datos del subsistema CLEP.



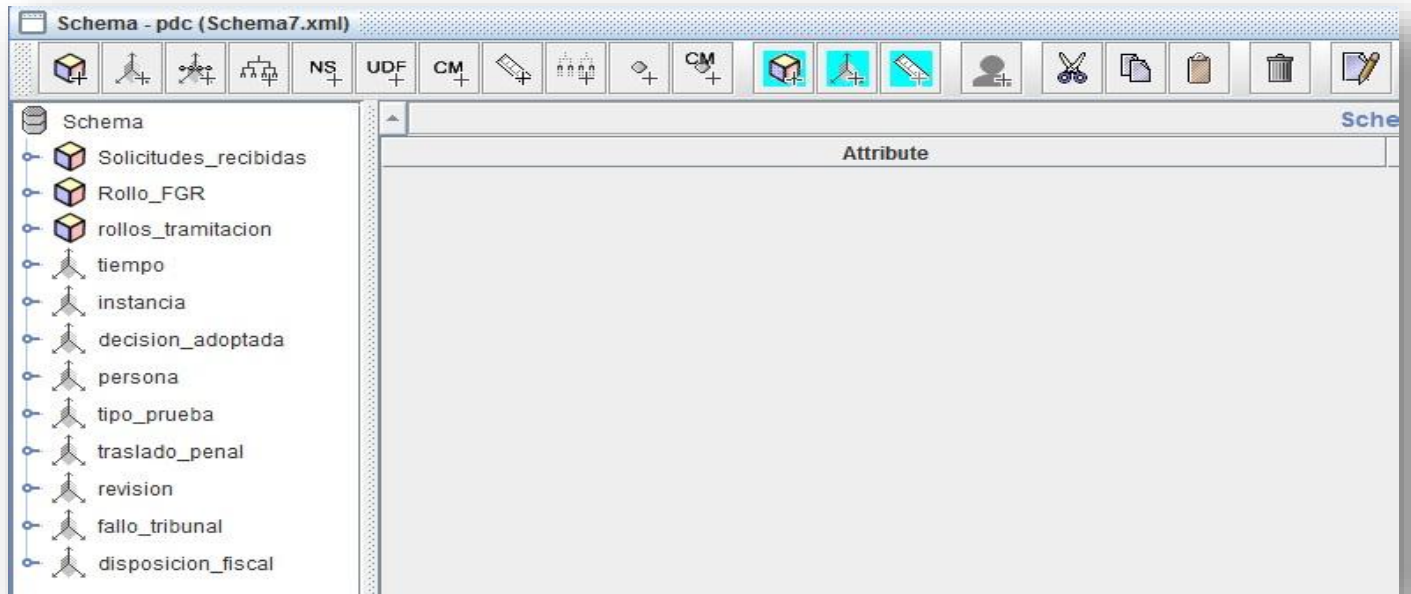
Anexo 10. Trabajo general para el Mercado de Datos del subsistema PDC.



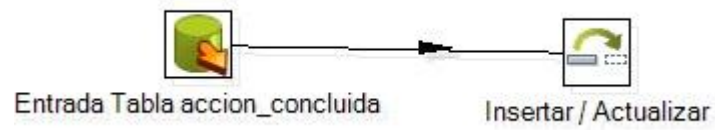
Anexo 11. Cubo de Datos para el Mercado de Datos del subsistema VF.



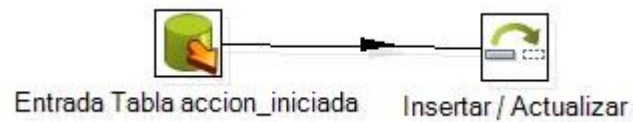
Anexo 12. Cubo de Datos para el Mercado de Datos del subsistema CLEP.



Anexo 13. Cubo de Datos para el Mercado de Datos del subsistema PDC.



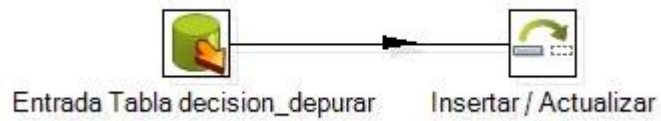
Anexo 14. Dimensión acciones concluidas.



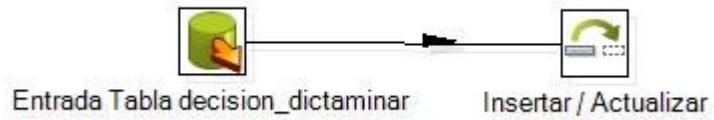
Anexo 15. Dimensión acciones iniciadas.



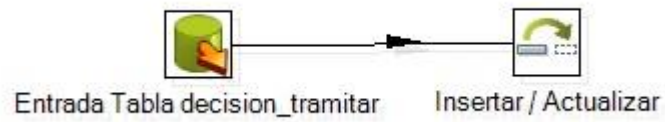
Anexo 16. Dimensión categoría.



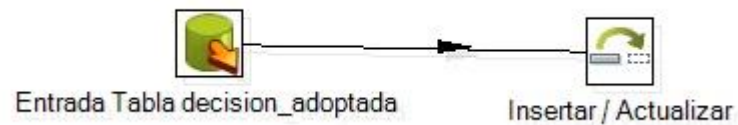
Anexo 17. Dimensión decisión depurar.



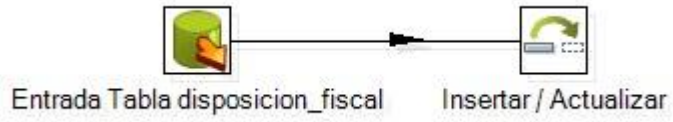
Anexo 18. Dimensión decisión dictaminar.



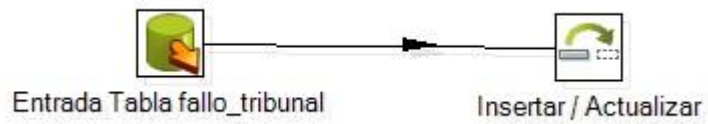
Anexo 19. Dimensión decisión tramitar.



Anexo 20. Dimensión decisión adoptada.



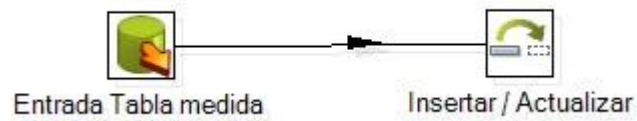
Anexo 21. Dimensión disposición fiscal.



Anexo 22. Dimensión fallo del tribunal.



Anexo 23. Dimensión instancia.



Anexo 24. Dimensión medida.



Anexo 25. Dimensión persona.



Anexo 26. Dimensión procedencia.



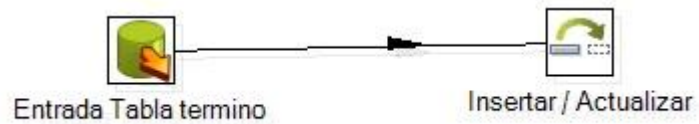
Anexo 27. Dimensión proroga.



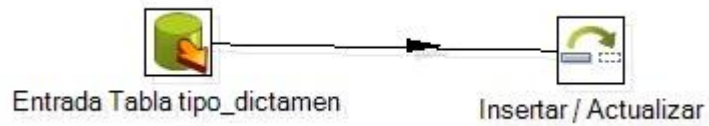
Anexo 28. Dimensión revisión



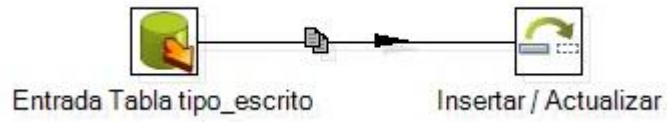
Anexo 29. Dimensión tema.



Anexo 30. Dimensión termino.



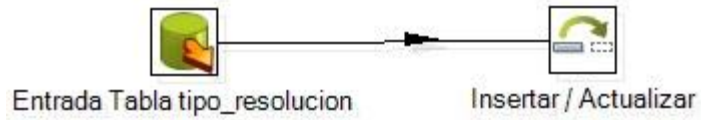
Anexo 31. Dimensión tipo de dictamen.



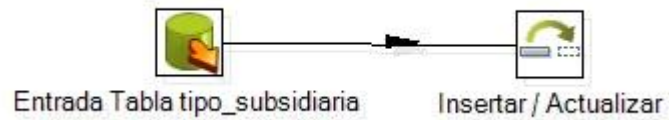
Anexo 32. Dimensión tipo de escrito.



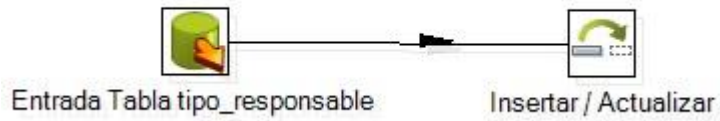
Anexo 33. Dimensión tipo de prueba.



Anexo 34. Dimensión tipo de resolución.



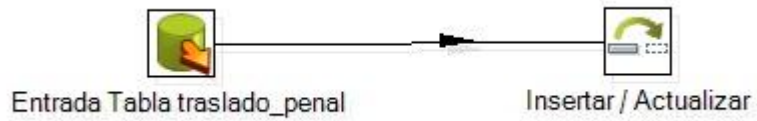
Anexo 35. Dimensión tipo de subsidiaria.



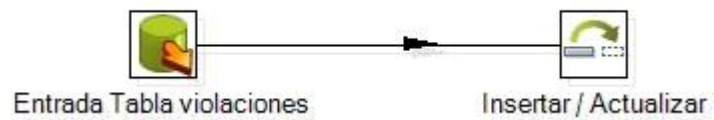
Anexo 36. Dimensión tipo de responsable.



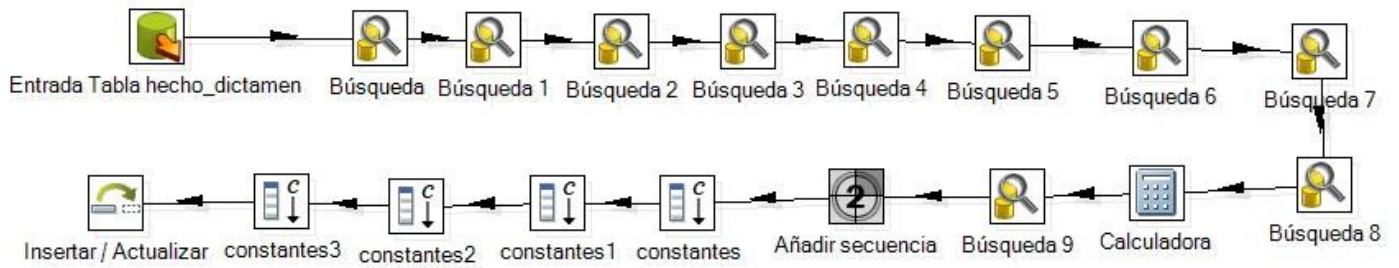
Anexo 37. Dimensión traslado órgano.



Anexo 38. Dimensión traslado penal.



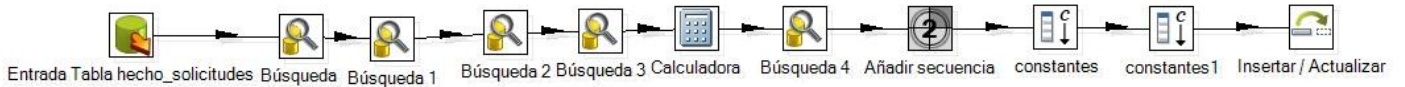
Anexo 39. Dimensión violaciones.



Anexo 40. Hecho dictamen.



Anexo 41. Hecho registro de escrito.



Anexo 42. Hecho solicitudes recibidas.



Anexo 43. Hecho hechos.



Anexo 44. Hecho rollo de tramitación.

```

SELECT
  drollodictamen.id_proceso,
  nprocedencia.id_procedencia,
  nprocedencia.descripcion,
  dproceso.fecha_creacion,
  dfiscalia.id_fiscalia,
  ntipofiscalia.descripcion,
  ntiposubsiaria.id_tipo_subsiaria,
  ntiposubsiaria.descripcion,
  ntipodictamen.id_tipo_dictamen,
  ntipodictamen.descripcion,
  ndecisiondictamen.id_decision_dictamen,
  ndecisiondictamen.descripcion,
  dpersona.id_persona,
  dpersona.raza,
  dpersona.sexo,
  nterminotiempo.id_termino,
  nterminotiempo.cant_dias,
  nfallotribunal.id_fallo,
  nfallotribunal.descripcion,
  ntiporesolucionjudicial.id_tipo_resolucion,
  ntiporesolucionjudicial.descripcion
FROM
  dictamen.drollodictamen
  inner join base.dproceso on drollodictamen.id_proceso = dproceso.id_proceso
  inner join dictamen.nprocedencia on drollodictamen.id_procedencia = nprocedencia.id_procedencia
  inner join base.dfiscalia on dfiscalia.id_fiscalia = dproceso.id_fiscalia
  inner join base.ntipofiscalia on ntipofiscalia.id_tipo_fiscalia = dfiscalia.id_tipo_fiscalia
  inner join dictamen.ntiposubsiaria on ntiposubsiaria.id_tipo_subsiaria =
drollodictamen.id_tipo_subsiaria
  inner join dictamen.ntipodictamen on ntipodictamen.id_tipo_dictamen =
drollodictamen.id_tipo_dictamen
  inner join base.ddocumento on ddocumento.id_proceso = dproceso.id_proceso

  left join dictamen.ddictamenclap on ddictamenclap.id_documento = ddocumento.id_documento

  inner join dictamen.ndecisiondictamen on ddictamenclap.id_decision_dictamen =
ndecisiondictamen.id_decision_dictamen
  inner join base.dpersonaproceso on dpersonaproceso.id_proceso = dproceso.id_proceso
  inner join base.dpersona on dpersonaproceso.id_persona = dpersona.id_persona
  inner join base.ntipoproceso on ntipoproceso.id_tipo_proceso = dproceso.id_tipo_proceso
  inner join base.nterminotiempo on nterminotiempo.id_tipo_proceso = ntipoproceso.id_tipo_proceso

  left join dictamen.drecursosuplica on drecursosuplica.id_documento = ddocumento.id_documento,

base.dresolucionjudicial,
base.ntiporesolucionjudicial,
base.nfallotribunal
Where
  (dictamen.ddictamenclap.id_resolucion_judicial = dresolucionjudicial.id_resolucion_judicial
or drecursosuplica.id_resolucion_judicial = dresolucionjudicial.id_resolucion_judicial or
dresolucionjudicial.id_resolucion_judicial is null) and
  (ntiporesolucionjudicial.id_tipo_resolucion = dresolucionjudicial.id_tipo_resolucion or
dresolucionjudicial.id_tipo_resolucion is null) and
  (nfallotribunal.id_fallo = dresolucionjudicial.id_fallo or dresolucionjudicial.id_fallo is
null)

```

Anexo 45. Consulta para el hecho dictámenes.

```
SELECT
ntemaclep.descripcion,
ntemaclep.id_tema_clep,
ddecisiondepurar.id_desicion_depurar,
ddecisiondepurar.discriminante,
dregistroescritoclep.id_proceso,
dpersona.id_persona,
dpersona.sexo,
dpersona.raza,
dproceso.fecha_creacion,
nfallotribunal.id_fallo,
nfallotribunal.descripcion,
dorganismotraslado.id_organismo_traslado,
norganismo.descripcion,
ntipoescrito.id_tipo_escrito,
ntipoescrito.descripcion,
ntipofiscalia.descripcion,
dfiscalia.id_fiscalia,
ddepuracion.no_radicacion,
ddepuracion.id_depurar

FROM
qpdclep.dregistroescritoclep,
qpdclep.ntemaclep,
qpdclep.dplanteamientoclep,
qpdclep.dtrasladodefinitivoclep,
base.ddepuracion,
base.ddecisiondepurar,
base.dorganismotraslado,
base.norganismo,
base.ntipoescrito,
base.dfiscalia,
base.ntipofiscalia,
base.dproceso,
base.dpersonaproceso,
base.dpersona,
base.nfallotribunal,
qpdclep.drespuestaimpugnacion
WHERE
dregistroescritoclep.id_depurar = ddepuracion.id_depurar AND
dregistroescritoclep.id_tipo_escrito = ntipoescrito.id_tipo_escrito AND
dregistroescritoclep.id_planteamiento = dplanteamientoclep.id_planteamiento AND
dregistroescritoclep.id_proceso = dproceso.id_proceso AND
dplanteamientoclep.id_tema_clep = ntemaclep.id_tema_clep AND
dtrasladodefinitivoclep.id_organismo_traslado = dorganismotraslado.id_organismo_traslado AND
ddepuracion.id_desicion_depurar = ddecisiondepurar.id_desicion_depurar AND
ddecisiondepurar.id_desicion_depurar = dtrasladodefinitivoclep.id_desicion_depurar AND
dorganismotraslado.id_traslado = norganismo.id_traslado AND
dfiscalia.id_tipo_fiscalia = ntipofiscalia.id_tipo_fiscalia AND
dproceso.id_fiscalia = dfiscalia.id_fiscalia AND
dpersonaproceso.id_proceso = dproceso.id_proceso AND
dpersonaproceso.id_persona = dpersona.id_persona AND
nfallotribunal.id_fallo = drespuestaimpugnacion.id_fallo AND
```

```
drespuestaimpugnacion.id_desicion_depurar = ddecisiondepurar.id_desicion_depurar;
```

Anexo 46. Consulta para el hecho registro de escrito.

```
SELECT
  ddecisiontramitar.discriminante,
  ddecisiontramitar.id_decision_tramitar,
  dprorroga.decision,
  dprorroga.id_documento,
  dfiscalia.id_fiscalia,
  ntipofiscalia.descripcion,
  nterminotiempo.id_termino,
  nterminotiempo.cant_dias,
  drolloquejaclep.id_proceso,
  dproceso.fecha_creacion
FROM
  qpdclep.drolloquejaclep,
  base.ddecisiontramitar,
  base.dprorroga,
  base.ddocumento,
  base.dproceso,
  base.dfiscalia,
  base.ntipofiscalia,
  base.ntipoproceso,
  base.nterminotiempo
WHERE
  ddecisiontramitar.id_decision_tramitar = drolloquejaclep.id_decision_tramitar AND
  ddocumento.id_documento = dprorroga.id_documento AND
  dproceso.id_proceso = drolloquejaclep.id_proceso AND
  dfiscalia.id_fiscalia = dproceso.id_fiscalia AND
  ntipofiscalia.id_tipo_fiscalia = dfiscalia.id_tipo_fiscalia AND
  ntipoproceso.id_tipo_proceso = dproceso.id_tipo_proceso AND
  nterminotiempo.id_tipo_proceso = ntipoproceso.id_tipo_proceso;
```

Anexo 47. Consulta para el hecho rollo de quejas.

```
SELECT
  dfiscalia.id_fiscalia,
  ntipofiscalia.descripcion,
  dproceso.fecha_creacion,
  dproceso.id_proceso,
  dpersona.id_persona,
  dpersona.sexo,
  ndecisionadoptada.id_decision_adop,
  ndecisionadoptada.descripcion,
  dtrasladopenal.id_trasladar_a,
  dtrasladopenal.descripcion
FROM
  penales.dsolicitud,
  base.dproceso,
  base.dfiscalia,
  base.ntipofiscalia,
  base.dpersona,
```



```

base.dpersonaproceso,
penales.ndecisionadoptada,
penales.ddepuracion,
penales.dtrasladopenal,
base.ddocumento
WHERE
dsolicitud.id_proceso = dproceso.id_proceso AND
dfiscalia.id_fiscalia = dproceso.id_fiscalia AND
ntipofiscalia.id_tipo_fiscalia = dfiscalia.id_tipo_fiscalia AND
dpersonaproceso.id_proceso = dproceso.id_proceso AND
dpersonaproceso.id_persona = dpersona.id_persona AND
ddepuracion.id_depuracion = dsolicitud.id_depuracion AND
ddepuracion.id_decision_adop = ndecisionadoptada.id_decision_adop AND
dtrasladopenal.id_documento = ddocumento.id_documento AND
ddocumento.id_proceso = dproceso.id_proceso;

```

Anexo 47. Consulta para el hecho solicitudes recibidas.

```

SELECT
dceduladeemplazamiento.solicito_prueba,
dceduladeemplazamiento.id_proceso,
dfiscalia.id_fiscalia,
ntipofiscalia.descripcion,
dproceso.fecha_creacion,
ntipoprueba.descripcion,
ntipoprueba.id_prueba
FROM
penales.dceduladeemplazamiento,
base.dfiscalia,
base.dproceso,
base.ntipofiscalia,
penales.ntipoprueba,
penales.dsolicitudprueba_ntipoprueba,
base.ddocumento
WHERE
dfiscalia.id_fiscalia = dproceso.id_fiscalia AND
dproceso.id_proceso = dceduladeemplazamiento.id_proceso AND
ntipofiscalia.id_tipo_fiscalia = dfiscalia.id_tipo_fiscalia AND
ntipoprueba.id_prueba = dsolicitudprueba_ntipoprueba.id_prueba AND
dsolicitudprueba_ntipoprueba.id_documento = ddocumento.id_documento AND
ddocumento.id_proceso = dproceso.id_proceso;

```

Anexo 48. Consulta para el hecho rollo de la FGR.

```

SELECT
ddecisiontramitar.discriminante,
ddecisiontramitar.id_decision_tramitar,
dprorroga.decision,
dprorroga.id_documento,
dfiscalia.id_fiscalia,
ntipofiscalia.descripcion,
nterminotiempo.id_termino,
nterminotiempo.cant_dias,

```

```

drolloquejaclep.id_proceso,
dproceso.fecha_creacion
FROM
qpdclep.drolloquejaclep,
base.ddecisiontramitar,
base.dprorroga,
base.ddocumento,
base.dproceso,
base.dfiscalia,
base.ntipofiscalia,
base.ntipoproceso,
base.terminotiempo
WHERE
ddecisiontramitar.id_decision_tramitar = drolloquejaclep.id_decision_tramitar AND
ddocumento.id_documento = dprorroga.id_documento AND
ddocumento.id_proceso = dproceso.id_proceso AND
dproceso.id_proceso = drolloquejaclep.id_proceso AND
dfiscalia.id_fiscalia = dproceso.id_fiscalia AND
ntipofiscalia.id_tipo_fiscalia = dfiscalia.id_tipo_fiscalia AND
ntipoproceso.id_tipo_proceso = dproceso.id_tipo_proceso AND
nterminotiempo.id_tipo_proceso = ntipoproceso.id_tipo_proceso;

```

Anexo 49. Consulta para el hecho rollo de tramitación.

```

SELECT
dproceso.id_proceso,
dproceso.fecha_creacion,
daccion.reinspeccion_planificada,
daccion.pertenece_sectorializada,
daccion.tipo_accion,
dprorrogavf.decision,
dprorrogavf.id_documento,
ntipofiscalia.descripcion,
dfiscalia.id_fiscalia,
dinicio.id_documento,
dconclusiones.id_documento,
dactareunion.hora,

(CASE

WHEN

daccion.pertenece_sectorializada = TRUE

THEN count(daccion.pertenece_sectorializada)

END) as cant_accion_sectorializada,
(CASE

WHEN

daccion.reinspeccion_planificada = TRUE

THEN count(daccion.reinspeccion_planificada)

```

```

END) as cant_reinspeccion_planificada,
(CASE

WHEN

daccion.tipo_accion = TRUE

THEN 1

ELSE

2

END) as cant_tipo_accion
FROM
vf.daccion,
base.dproceso,
vf.dprorrogavf,
base.ddocumento,
base.dfiscalia,
base.ntipofiscalia,
vf.dinicio,
vf.dconclusiones,
vf.dactareunion

WHERE
dproceso.id_proceso = daccion.id_proceso AND
dprorrogavf.id_documento = ddocumento.id_documento AND
ddocumento.id_proceso = dproceso.id_proceso AND
dfiscalia.id_fiscalia = dproceso.id_fiscalia AND
ntipofiscalia.id_tipo_fiscalia = dfiscalia.id_tipo_fiscalia AND
dinicio.id_documento = dactareunion.id_documento AND
dconclusiones.id_documento = dactareunion.id_documento AND
dactareunion.id_documento = ddocumento.id_documento

GROUP BY
dproceso.id_proceso,daccion.reinspeccion_planificada,daccion.pertenece_sectorializada,daccion.ti
po_accion,dprorrogavf.decision,dprorrogavf.id_documento,ntipofiscalia.descripcion,dfiscalia.id_f
iscalia,dinicio.id_documento,dconclusiones.id_documento,dactareunion.hora,dactareunion.id_docume
nto;

```

Anexo 50. Consulta para el hecho acciones.

```

SELECT
dhecho.id_proceso,
dproceso.fecha_creacion,
daccion.pertenece_sectorializada,
daccion.tipo_accion,
dmedida.id_med_solicitada,
dmedida.disciplinaria,
dviolacion.id_violacion,
dviolacion.causas_condiciones,
ntiporesponsable.descripcion,
ntiporesponsable.id_tipo_responsable,
ntipofiscalia.descripcion,

```

```

dfiscalia.id_fiscalia,
norganismo.descripcion,
dorganismotraslado.id_organismo_traslado,
dcategoria.discriminante,
dcategoria.id_categoria,
nterminotiempo.id_termino,
nterminotiempo.cant_dias,
(CASE

WHEN

daccion.pertenece_sectorializada = TRUE

THEN count(daccion.pertenece_sectorializada)

END) as cant_accion_sectorializada,
(CASE

WHEN

daccion.tipo_accion = TRUE

THEN 1

ELSE

2

END) as cant_tipo_accion
FROM
vf.dmedida,
vf.dhecho,
vf.dsolicitada,
vf.daccion,
base.dproceso,
vf.dviolacion,
vf.ntiporesponsable,
vf.ddatosresponsable,
base.dfiscalia,
base.dorganismotraslado,
base.ntipofiscalia,
base.norganismo,
vf.dcategoria,
vf.dinformedenunciahechosdelictivos,
base.ddocumento,
base.nterminotiempo,
base.ntipoproceso
WHERE
dmedida.id_proceso = dhecho.id_proceso AND
dhecho.id_proceso = dproceso.id_proceso AND
dsolicitada.id_med_solicitada = dmedida.id_med_solicitada AND
daccion.id_proceso = dhecho.id_proceso_accion AND
dviolacion.id_hecho = dhecho.id_proceso AND
ddatosresponsable.id_medida = dmedida.id_medida AND
ddatosresponsable.id_tipo_responsable = ntiporesponsable.id_tipo_responsable AND
dfiscalia.id_fiscalia = dproceso.id_fiscalia AND

```

```

ntipofiscalia.id_tipo_fiscalia = dfiscalia.id_tipo_fiscalia AND
norganismo.id_traslado = dorganismotraslado.id_traslado AND
dcategoria.id_categoria = dviolacion.id_categoria AND
dinformedenunciahechosdelictivos.id_documento = ddocumento.id_documento AND
dinformedenunciahechosdelictivos.id_organismo_traslado =
dorganismotraslado.id_organismo_traslado AND
ddocumento.id_proceso = dhecho.id_proceso AND
nterminotiempo.id_tipo_proceso = ntipoproceso.id_tipo_proceso AND
ntipoproceso.id_tipo_proceso = dproceso.id_tipo_proceso
GROUP BY
dhecho.id_proceso, dproceso.id_proceso, daccion.pertenece_sectorializada, daccion.tipo_accion, ntipo
fiscalia.descripcion, dfiscalia.id_fiscalia, nterminotiempo.id_termino, dmedida.id_med_solicitada, d
medida.disciplinaria, dviolacion.id_violacion, dviolacion.causas_condiciones, ntiporesponsable.desc
ripcion, ntiporesponsable.id_tipo_responsable, dcategoria.discriminante, norganismo.descripcion, dor
ganismotraslado.id_organismo_traslado, dcategoria.id_categoria;

```

Anexo 51. Consulta para el hecho hechos.

```

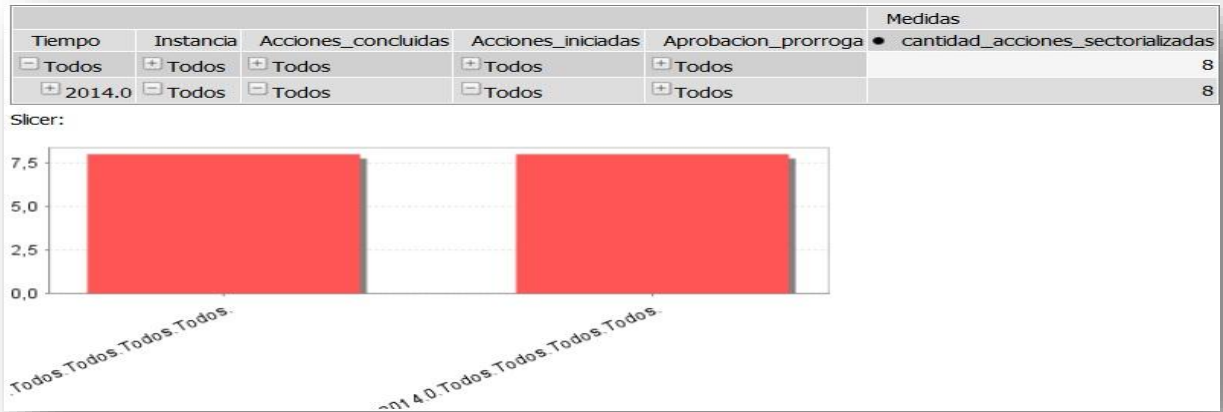
SELECT
dfiscalia.id_fiscalia,
ntipofiscalia.descripcion,
dproceso.id_proceso,
dproceso.fecha_creacion,
ddisposicionfiscal.decision_fiscal,
ddisposicionfiscal.id_disposicion,
nfallotribunal.id_fallo,
nfallotribunal.descripcion,
drevision.id_revision,
ndecision.descripcion
FROM
penales.drollopenal,
base.dproceso,
base.dfiscalia,
base.ntipofiscalia,
penales.ddisposicionfiscal,
base.dresolucionjudicial,
base.nfallotribunal,
penales.drevision,
penales.ndecision,
penales.descritopromocion,
base.ddocumento
WHERE
drollopenal.id_proceso = dproceso.id_proceso AND
dfiscalia.id_fiscalia = dproceso.id_fiscalia AND
dfiscalia.id_tipo_fiscalia = ntipofiscalia.id_tipo_fiscalia AND
ddisposicionfiscal.id_disposicion = drollopenal.id_disposicion AND
dresolucionjudicial.id_resolucion_judicial = drollopenal.id_resolucion_judicial AND
nfallotribunal.id_fallo = dresolucionjudicial.id_fallo AND
drevision.id_decision = ndecision.id_decision AND
drevision.id_revision = descritopromocion.id_revision AND
descritopromocion.id_documento = ddocumento.id_documento AND
ddocumento.id_proceso = dproceso.id_proceso;

```

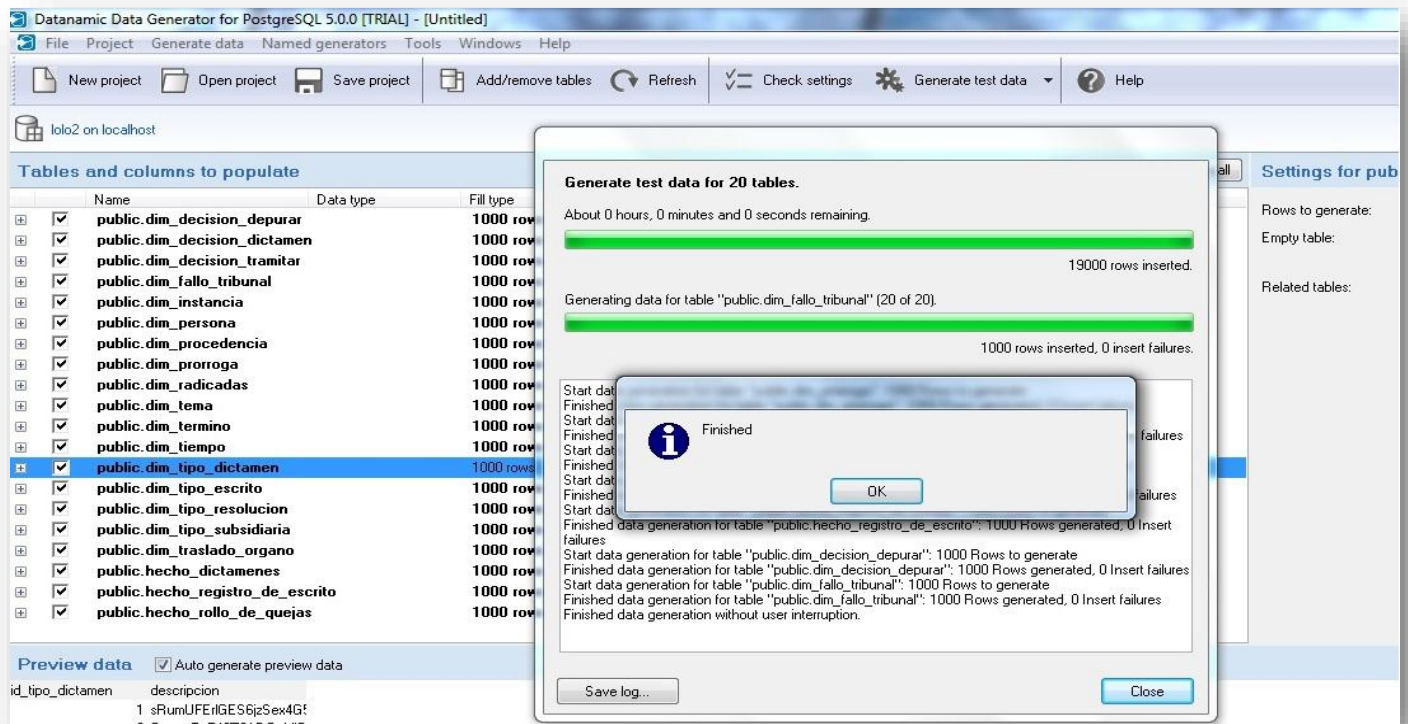
Anexo 52. Consulta para el hecho rollo de tramitación.



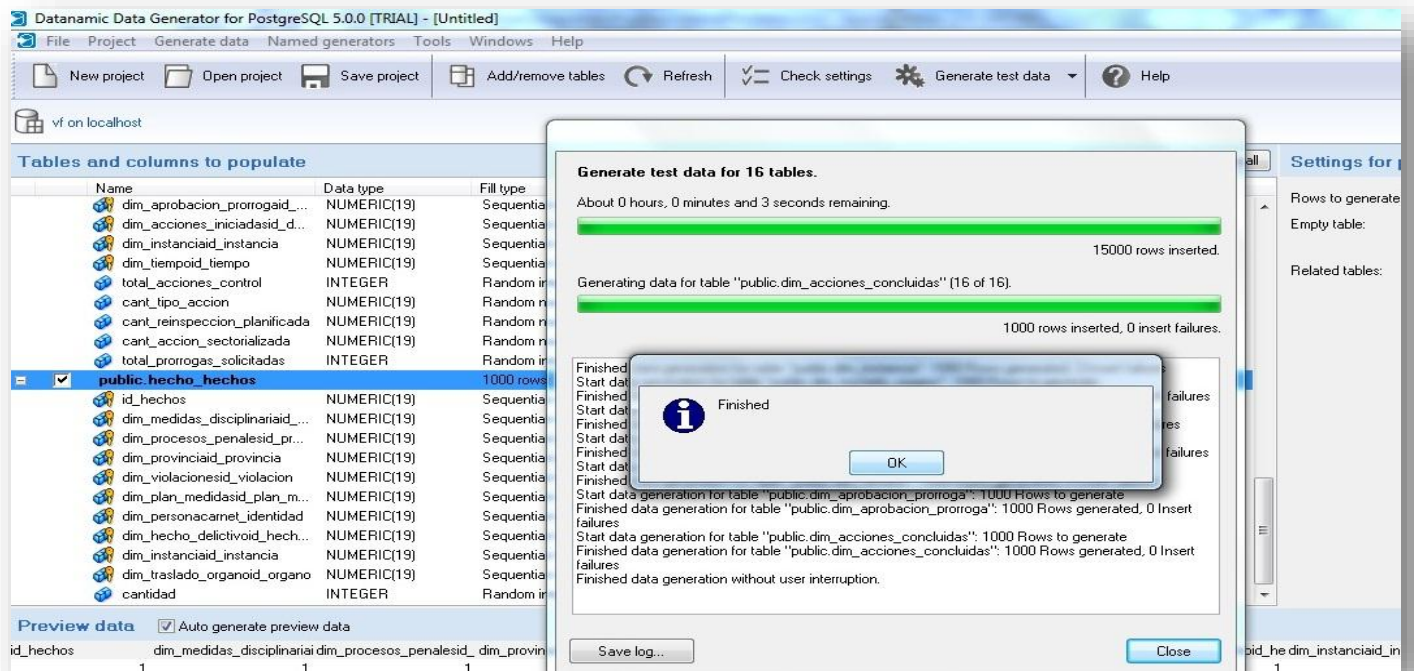
Anexo 53. Representación en el navegador del cubo Rollo_de_queja.



Anexo 54. Representación en el navegador del cubo Acciones.



Anexo 55. Resultado de la prueba de Carga en Datanamic Data Generator para el mercado del subsistema CLEP.



Anexo 56. Resultado de la prueba de Carga en Datanamic Data Generator para el mercado del subsistema VF.