

Universidad de las Ciencias Informáticas

Facultad 1



Título: Solución de Inteligencia de Negocios para el Sistema Único de Identificación Nacional.

Trabajo de Diploma para optar por el título de Ingeniero en Ciencias Informáticas.

Autor(es): Adonis Gonzáles Santos.

Yunier Arcelio Pérez Tamayo.

Tutor: MsC. Erik de la Vega García.

La arcilla fundamental de nuestra obra es la juventud; en ella depositamos nuestra esperanza y la preparamos para tomar de nuestras manos la bandera.

Ernesto Che Guevara

Declaramos ser autores del presente trabajo de diploma y reconocemos a la Universidad de las Ciencias Informáticas los derechos patrimoniales de la misma, con carácter exclusivo. Autorizamos a dicho centro para que haga el uso que estime pertinente con este trabajo.

Para que así conste firmamos la presente a los ____ días del mes de _____ del año _____.

Yunier Arcelio Pérez Tamayo

Adonis González Santos

Firma del Autor

Firma del Autor

Msc. Erik de la Vega García

Firma del Tutor

Dedicatoria

A mi familia por el apoyo brindado en el transcurso de todos estos años de estudio,

En especial a mi madre.

Adonis

A mi familia que me ha apoyado tanto para que cumpla todas mis metas,

En especial a mi madre que lo ha dado todo por mí.

Yunier

A mi familia, mi mamá, mis abuelos, mi tío, mi tía, mi papá, mi hermanito, a toda mi familia por todo el apoyo que me han brindado.

A mi novia Arianna por estar a mi lado durante estos años de universidad en los momentos de alegría, en los de tristeza.

A mis compañeros, Sandy, Tony, Yosvany, Valladares, Dianela, Dainer, Yunier, José que me han brindado su ayuda durante estos años de estudio.

A mis compañeros de ocio o como le decimos de manera jovial: la comunidad; por hacer de todos este tiempo más llevadero, más alegre.

Al tutor del trabajo de diploma Erik y a Yudenia por la ayuda brinda.

Por ultimo pero no menos importante al Comandante en Jefe por haber creado esta Universidad que me ha formado, me ha preparado no solo como ingeniero sino que ha hecho de mi mejor persona.

Adonis

A mi familia, mi mamá, mi papá, mis hermanas, mis tías, mis primos por todo el apoyo que me han dado a lo largo de estos años de universidad.

A mis amigos, Dianela, Evelyn, Adonis, Yosvany, Tony, José, Sandy, Yicel, Karelia, Yisel, Yannier por haber sido mi familia aquí en la universidad.

A mis compañeros y amigos de la comunidad de Firefoxmania, Erick, Yunier, Diego, Roberto, Manuel, Abelardo, Jacobo por haber formado parte de este gran logro de colaboración e intercambio.

A nuestro tutor Erik y a nuestra jefa de proyecto Yudenia por la ayuda brindada durante la realización de este trabajo.

Por último a esta Revolución por permitir que todos los jóvenes de este país puedan estudiar y formarse como profesionales y posteriormente contribuir a defender las conquistas obtenidas a lo largo de estos años.

Yunier

Resumen

Resumen

Cuba se encuentra envuelta en un proceso de informatización que comprende varias de las esferas de la sociedad y la economía, donde el Ministerio del Interior (MININT) no está exento; para lograr un mejor servicio se dispone a realizar la automatización de las áreas de la Dirección de Identificación y Registro (DIR), las cuales tienen como principios llevar a cabo un control estricto de la identificación de la población nacional.

Con este objetivo se decide crear el Sistema de Único de Identificación Nacional (SUIN), desarrollado por el Centro de Identificación y Seguridad Digital (CISED) en conjunto con el MININT.

A través de dicho sistema se realizarán más de 6000 trámites diarios, generando un gran cúmulo de información, por lo que se hace imprescindible recolectar y procesar la misma, ofreciendo a los funcionarios, ministerios (a los cuales se les brindan reportes sobre diversos trámites) y altos ejecutivos acceso a la información clave acerca de las diversas operaciones y procesos para facilitar la gestión histórica y la generación de reportes, facilitando su posterior análisis para la toma de decisiones.

El presente trabajo de diploma tiene como objetivo general: desarrollar una solución de inteligencia de negocios para el Sistema Único de Identificación Nacional mejorando así la gestión histórica de la información y la generación de reportes sirviendo de apoyo a la toma de decisiones.

Esta solución permite a los funcionarios del MININT recolectar y procesar la información generada a partir de los trámites realizados en la DIR facilitando la gestión histórica, proporcionando acceso a información clave para la toma de decisiones.

Palabras claves:

DIR, gestión histórica, inteligencia de negocios, toma de decisiones.

Índice de contenido

Índice

Resumen	VI
Índice	VII
Introducción	1
Capítulo 1. Fundamentación teórica	5
1.1. Introducción	5
1.2. Inteligencia de Negocios	5
1.3. Almacenamiento de datos	7
1.4. Metodología de desarrollo	15
1.5. Herramienta de diseño y modelado	19
1.6. Herramientas para la realización de Inteligencia de Negocios	19
1.7. Herramientas para la realización del proceso de ETL	21
1.8. Sistema Gestor de Base de Datos	22
1.9. Herramienta para la Gestión de la Base de Datos	23
1.10. Conclusiones.....	23
Capítulo 2. Análisis y diseño de la solución	25
2.1. Introducción	25
2.2. Descripción del negocio.....	25
2.3. Necesidades del usuario.....	26
2.4. Reglas de Negocio	27
2.5. Especificación de requisitos.....	28
2.6. Modelo de arquitectura a seguir.....	32
2.7. Modelo de datos Dimensional.....	35
2.8. Esquema de Seguridad	38
2.9. Conclusiones	40
Capítulo 3. Implementación de la solución.....	41
3.1. Introducción	41
3.2. Implementación de la base de datos.....	41
3.3. Implementación de los procesos de BI	48
3.4. Diagrama de despliegue	52

Índice de contenido

3.5. Conclusiones	53
Capítulo 4. Validación y pruebas de la Solución de Inteligencia de Negocios	54
4.1. Introducción	54
4.2. Validación y pruebas	54
4.3. Tipos de pruebas a realizar	54
4.4. Resultados de la aplicación de las pruebas	55
4.5. Análisis de la rapidez de obtención de los reportes	59
4.6. Beneficios de la aplicación.....	61
4.7. Conclusiones	61
Conclusiones	62
Trabajos citados	63
Anexos.....	66
Anexo 1. Descripción de los requisitos funcionales	66
Anexo 2. Modelo de dominio del DGM	74
Anexo 3. Modelos de datos del <i>data warehouse</i>	75
Anexo 4. ETL's realizadas	76
Anexo 5. Reportes implementados.....	86

Índice de figuras

Figura 1. Niveles de las soluciones inteligencia de negocio.	6
Figura 2. Estructura de un cubo OLAP.....	13
Figura 3. Ciclo de vida que plantea la metodología Kimball.	16
Figura 4. Pirámide informativa.	26
Figura 5. Componentes que conforman el <i>data warehouse</i>	32
Figura 6. Diagrama de flujo del proceso de integración de los datos.	34
Figura 7. Modelo de datos.	38
Figura 8. ETL para la carga de los cambios de dirección.....	45
Figura 9. ETL para la carga de los datos relativos a los trámites.	46
Figura 10. Job para la carga de los datos de las personas.	47
Figura 11. Job para la carga de los datos de los trámites.	48
Figura 12. Cubos OLAP.....	49
Figura 13. Mapa de navegación.....	50
Figura 14. Reporte cantidad por tipo de trámite.	51
Figura 15. Cantidad de trámites por actualización y solicitud de documentos.....	51
Figura 16. Coeficiente de fiabilidad.	52
Figura 17. Modelo de despliegue.....	52
Figura 18. Por ciento de errores por iteración.	55
Figura 19. Datos que se encuentran persistidos en la base de datos.	56
Figura 20. Reporte mostrado en la aplicación.....	56
Figura 21. Muestra el tiempo medio de respuesta por la cantidad de usuarios.	57
Figura 22. Carta de aceptación.....	59
Figura 23. Modelo de dominio DGM.	74
Figura 24. Modelo de datos del <i>data warehouse</i>	75
Figura 25. ETL para la carga de los datos del DGM.....	76
Figura 26. ETL para la carga de los datos hacia DimDispositivo.....	76
Figura 27. ETL para la carga de los datos hacia DimEstacion.	76
Figura 28. ETL para la carga de los datos hacia DimOficina.	77
Figura 29. ETL para la carga de los datos hacia FactDispositivo.	77
Figura 30. ETL para la realización de la carga de los datos referentes a los cambios de dirección.....	78
Figura 31. ETL que muestra la carga de los datos hacia la dimensión DimDirActual.	79
Figura 32. ETL que muestra la carga de los datos hacia la dimensión DimDirProcedencia.	80
Figura 33. ETL que muestra la carga de los datos hacia la tabla hecho FactCambioDir.	80
Figura 34. ETL que realiza la carga de los datos referentes a las persona.	81
Figura 35. ETL que muestra la carga de los datos hacia la dimensión DimPersona.	81
Figura 36. ETL que muestra la carga de los datos hacia la dimensión DimUbicacion.	82
Figura 37. ETL que muestra la carga de los datos hacia la tabla hecho FactPersona.....	82
Figura 38. ETL que realiza la carga de los datos referentes a los trámites.	83
Figura 39. ETL que muestra la carga de los datos hacia la tabla hecho FactTramitesOficina.....	84

Índice de figuras

Figura 40. ETL que muestra la carga de los datos hacia la tabla hecho FactTramite.....	85
Figura 41. Reporte referente al tiempo promedio de servicio.	86
Figura 42. Reporte referente al coeficiente de fiabilidad.	87
Figura 43. Reporte referente a los trámites de actualización y solicitud de documentos.	87
Figura 44. Reporte de trámites por tipo.....	88

Índice de tablas

Tabla 1. Comparación entre los modos de almacenamiento.....	10
Tabla 2. Matriz Bus.....	37
Tabla 3. Seguridad en la base de datos.....	39
Tabla 4. Seguridad en la aplicación.	39
Tabla 5. Hechos y Dimensiones.....	41
Tabla 6. Secuencias	42
Tabla 7. Índices.	43
Tabla 8. Tiempo de obtención de los reportes del SUIN.	60
Tabla 9. Tiempo de obtención de los reportes por la solución implementada.....	60

Introducción

Introducción

En el mundo actual los volúmenes de información que se manejan en el ámbito industrial, en empresas y gobiernos son enormes; así como los flujos de trabajo que se generan día a día son más elevados y complejos. Típicamente, dicha información se almacena a través de sistemas transaccionales en bases de datos relacionales que cumplen la función de almacenaje y administración de los datos. Pero la idea es que dejen de ser simples datos para convertirse en información que enriquezca las decisiones de los ejecutivos. Provocando que con frecuencia, el proceso de transformación y análisis de la información que se genera se torne más compleja y costosa, aquí es donde entra a jugar un papel fundamental en la toma de decisiones las soluciones de inteligencia de negocios (BI, por sus siglas en inglés).

Las tecnologías de BI intentan revertir esta situación ayudando a recolectar, comprender mejor y más rápido los datos, mediante el examen de los datos acumulados generando escenarios, pronósticos y reportes; posibilitan un análisis conciso en el momento y lugar correcto, logrando mejorar la toma de decisiones, traduciéndose en la práctica como identificar y evaluar oportunidades o deficiencias implícitas en los procesos de negocio.

El concepto de BI se puede definir como el proceso de analizar los bienes o datos acumulados en la organización o empresa y extraer una cierta inteligencia o conocimiento de ellos. Dentro de la categoría de bienes se incluyen las bases de datos de clientes, información de la cadena de suministro, ventas personales y cualquier actividad de marketing o fuente de información relevante. (1)

Cuba se encuentra envuelta en un proceso de informatización que comprende varias de las esferas de la sociedad y la economía, donde el Ministerio del Interior (MININT) no está exento; para lograr un mejor servicio se dispone a realizar la automatización de las áreas de la Dirección de Identificación y Registro (DIR), las cuales tienen como principios llevar a cabo un control estricto de la identificación de la población nacional. Por la relevancia de esta dirección se hace necesario incorporar una mayor rapidez y facilidad en la realización de los procesos que se llevan a cabo. (2)

Con este objetivo se decide crear el Sistema de Único de Identificación Nacional (SUIN), desarrollado por el Centro de Identificación y Seguridad Digital (CISED) en conjunto con el MININT.

Dicho sistema se encontrará desplegado por las más de 150 oficinas pertenecientes a la DIR donde se realizan más de 6000 trámites diarios generando un crecimiento mensual de la base de datos de 3Gb (3),

Introducción

como resultado de guardar cada una de las faces por la que pasa el trámite, así como datos referentes a los dispositivos usados y los funcionarios que lo realizan; la información es almacenada en una base de datos relacional usando como gestor de base de datos Oracle 11gR2, donde la información se encuentre disgregada en alrededor de 255 tablas, producto del proceso de normalización, siendo optimizada para sistemas de procesamiento de transacciones en línea (OLTP por sus siglas en inglés), derivándose que la obtención de reportes tarde un promedio de 6 minutos la obtención de reportes lo cual no cumple con las necesidades de inmediatez de acceso a la información por parte de los funcionarios del MININT, provocando a su vez una disminución del rendimiento de la base de datos, donde por ejemplo el uso de CPU sobrepasa el 60%.

Debido a lo antes planteado se hace imprescindible recolectar y procesar el gran cúmulo de información generada por el sistema, ofreciendo a los funcionarios, ministerios (a los cuales se les brinda información sobre diversos trámites) y altos ejecutivos acceso a la información clave acerca de las diversas operaciones y procesos para facilitar la gestión histórica y la generación de reportes, facilitando su posterior análisis para la toma de decisiones, por lo que se plantea el siguiente problema científico: ¿Cómo mejorar la gestión de la información histórica y generación de reportes para el apoyo en la toma de decisiones del Sistema de Único de Identificación Nacional?

A partir del problema planteado constituye objeto de estudio la gestión de la información histórica y generación de reportes enmarcando su campo de acción las soluciones de inteligencia de negocios para el apoyo en la toma de decisiones del Sistema Único de Identificación Nacional (SUIN).

Como objetivo general se define: Desarrollar una solución de inteligencia de negocios para el Sistema Único de Identificación Nacional mejorando así la gestión histórica de la información y la generación de reportes sirviendo de apoyo a la toma de decisiones.

Se propone como hipótesis: Si se implementa una solución de inteligencia de negocios para el SUIN se logrará mejorar el acceso a la información de forma ágil para el apoyo a la toma de decisiones de dicho sistema.

Para dar cumplimiento a dicho objetivo general se trazan las siguientes tareas investigativas:

1. Estudio de los principales conceptos asociados a las soluciones de inteligencia de negocios.

Introducción

2. Estudio bibliográfico para la definición de la metodología y herramientas a utilizar.
3. Identificación y descripción de los requisitos de la solución de inteligencia de negocios para el SUIN.
4. Identificación y descripción de las medidas, dimensiones y hechos.
5. Desarrollo de la matriz dimensional.
6. Estructuración del modelo de datos.
7. Definición de la arquitectura de integración de los datos.
8. Definición de la arquitectura de información.
9. Diseño de los procesos de integración de los datos.
10. Implementación del modelo de datos.
11. Implementación de los flujos de las transformaciones y los trabajos.
12. Implementación de los cubos multidimensionales.
13. Implementación de los reportes candidatos.
14. Implementación de los niveles de acceso de los usuarios.
15. Validación de la solución obtenida.

Para asistir las tareas de la investigación se emplearán los siguientes métodos científicos:

Métodos teóricos:

- **Histórico-Lógico:** Se utilizó para hacer un análisis de la utilización de soluciones de inteligencia de negocios, así como el impacto de estas en las organizaciones.
- **Analítico-Sintético:** Se analiza toda la información recopilada a través de los diferentes medios bibliográficos que puedan servir para desarrollar mejor el sistema antes planteado.
- **Modelación:** Este método permitió la creación de modelos (propuestas, alternativas y estrategias) que visualizan una reproducción simplificada de la realidad y que consisten en descubrir y estudiar nuevas relaciones y cualidades del objeto de estudio.

Métodos empíricos:

- **Entrevista:** Se manifestó a través de la realización de entrevistas a los proveedores de requisitos, además es necesario el intercambio con el personal capacitado para obtener información referente

Introducción

al tema, criterios y corregir malas concepciones para un mayor entendimiento del problema a resolver y determinar las necesidades del cliente.

Como **posibles resultados** de esta investigación se espera:

- Obtener una versión estable de la solución de inteligencia de negocios para el SUIN.
- Brindar un documento con la especificación de requisitos del sistema desarrollado.
- Manual de Usuario de la aplicación.

El presente documento consta de cuatro capítulos, desarrollados a partir del estudio realizado de la información bibliográfica e institucional existente. La descripción de los mismos se presenta a continuación:

- **Capítulo 1. Fundamentación teórica:** En este capítulo se abordará toda la fundamentación teórica sobre el objeto de estudio, los principales conceptos, metodologías y herramientas para el desarrollo de la solución de inteligencia de negocios para el SUIN.
- **Capítulo 2. Análisis y diseño de la solución:** En este capítulo se realiza el análisis y diseño. Se definen los requisitos funcionales, no funcionales y de información, el modelo de datos, entre otros artefactos que ayudan a entender las necesidades de información.
- **Capítulo 3. Implementación de la solución:** En este capítulo se realizarán todas las transformaciones necesarias como Extraer, Transformar y Cargar (ETL), proceso que deja toda la información lista para la fase siguiente, Inteligencia del Negocios; donde se realizará la implementación del modelo de datos, los cubos OLAP, los reportes candidatos y la política de seguridad de los usuarios.
- **Capítulo 4. Validación y pruebas:** En este capítulo se realiza la validación y pruebas de la solución propuesta a través de casos de prueba y listas de chequeo, para así evaluar los resultados obtenidos.

Capítulo 1. Fundamentación teórica

Capítulo 1. Fundamentación teórica

1.1. Introducción

En el presente capítulo se definen y analizan los conceptos o elementos relacionados con los sistemas de inteligencia empresarial posibilitando una mejor comprensión de las características de estos sistemas y una valoración posibles vías de solución a la problemática planteada. También se realiza un estudio de las metodologías existentes con el objetivo de seleccionar la más viable, además de fundamentar las principales herramientas integradas utilizadas a nivel mundial para la implantación de soluciones de inteligencia de negocios para el desarrollo de la solución propuesta.

1.2. Inteligencia de Negocios

De acuerdo al *Data Warehousing Institute*, la definición de Inteligencia de Negocios es la siguiente: Unifica los datos, tecnología, análisis, y el conocimiento humano para optimizar las decisiones empresariales y, finalmente, impulsar el éxito de una empresa. Programas de BI suelen combinar un *data warehouse* empresarial y una plataforma de BI o conjunto de herramientas para transformar datos en información útil, empresarial procesable. (4)

La Inteligencia de Negocios es el proceso de analizar los datos acumulados en la empresa y extraer un conocimiento de ellos. Se incluyen las bases de datos de clientes, información de la cadena de suministro, ventas personales y cualquier actividad de *marketing* o fuente de información relevante para la empresa. (5)

Las soluciones de BI proporcionan amplias ventajas ya sean competitivas para las empresas o de índole cognoscitivo en temas no empresariales, al permitir que los datos se conviertan en un centro de beneficios, facilitando el análisis de información útil a las organizaciones para la toma de decisiones, entre ellas se encuentran:

- Ofrece a los empleados, socios y altos ejecutivos acceso a la información clave que necesitan para realizar sus tareas del día a día, y principalmente para poder tomar decisiones basadas en datos correctos y certeros.
- Mejora de la colaboración y la calidad de las decisiones, facilitando el acceso a la información en todos los niveles de la organización.
- Ayuda a proveer oportunamente y acertadamente acceso a la información acerca de cualquier servicio o producto, para sus usuarios finales.

Capítulo 1. Fundamentación teórica

- Crea nuevas estructuras de relaciones entre las empresas, sus clientes y asociados, la conectividad de personas, organizaciones y países, y el valor del conocimiento residente en la empresa.
- Su conocimiento y habilidades y el uso de sistemas inteligentes para la toma de decisiones.
- Simplifican muchas operaciones y procesos en el negocio.

La Inteligencia de Negocios se encuentra dividida en tres niveles atendiendo a quien va enfocada la información que se brinda, estos son:

Nivel Operativo: Permite que los empleados que trabajan con información operativa puedan recibir la misma de una manera oportuna, exacta y adecuada y se componen básicamente de herramientas de reportes u hojas de cálculo con un formato fijo cuya información se actualiza frecuentemente. (6)

Nivel Táctico: Permite que los analistas de datos y la gerencia media de la empresa utilicen herramientas de análisis y consulta con el propósito de tener acceso a la información sin intervención de terceros. (6)

Nivel Estratégico: Permite que la alta dirección de las empresas pueda analizar y monitorear tendencias, patrones, metas y objetivos estratégicos de la organización. (6)



Figura 1. Niveles de las soluciones inteligencia de negocio.

La solución de inteligencia que se debe implementar debe llevarse al nivel estratégico para los directivos del MININT mostrándoles el comportamiento de los indicadores organizacionales de sus áreas a través de

Capítulo 1. Fundamentación teórica

cuadros de mando permitiéndoles analizar y monitorear tendencias y a un nivel operativo para los distintos funcionarios para la obtención de reportes de manera oportuna.

1.3. Almacenamiento de datos

Desde el surgimiento de las tecnologías de almacenamiento de datos, estas se han convertido en una herramienta fundamental para el control y manejo de operaciones. Estas no solo son útiles para el manejo de operaciones y para almacenar información, sino que también nos ayudan en los procesos de toma de decisiones. A continuación se abordará sobre varias tecnologías de almacenamiento de datos que han surgido a través de los años y que forman hoy día un elemento importante para el almacenamiento de información en disímiles empresas.

1.3.1. Bases de Datos

Se puede definir una base de datos como un conjunto de datos estructurados, fiables y homogéneos, organizados independientemente en máquina, accesibles a tiempo real, compartibles por usuarios concurrentes que tienen necesidades de información diferente y no predecible en el tiempo. (7)

Existen varios tipos de bases de datos según la variabilidad de los datos almacenados:

Bases de datos estáticas: Son bases de datos de solo lectura, utilizadas primordialmente para almacenar datos históricos que posteriormente se pueden utilizar para estudiar el comportamiento de un conjunto de datos a través del tiempo, realizar proyecciones y tomar decisiones. (8)

Bases de datos dinámicas: Son bases de datos donde la información almacenada se modifica con el tiempo, permitiendo operaciones como actualización y adición de datos, además de las operaciones fundamentales de consulta. Un ejemplo de esto puede ser la base de datos utilizada en un sistema de información de una tienda, una farmacia, un videoclub, etc. (8)

Se propone el uso de una base de datos estática debido a que los datos luego de ser cargados, no son modificables facilitando la gestión histórica de la información.

1.3.2. Modos de almacenamiento

Existen tres modelos para el Proceso Analítico en Línea (OLAP, por sus siglas en inglés) de la información: Proceso Analítico en Línea Relacional (ROLAP, por sus siglas en inglés), Proceso Analítico en Línea Multidimensional (MOLAP, por sus siglas en inglés) y Proceso Analítico en Línea Híbrido (HOLAP, por sus siglas en inglés). El proceso de análisis se realiza de igual forma lo que varía en uno y otro caso es la metodología de almacenamiento. La forma de almacenamiento es crítica para garantizar la

Capítulo 1. Fundamentación teórica

velocidad de recuperación de la información, las zonas de ubicación de las agregaciones y el procesamiento de los datos en general.

Proceso Analítico Relacional ROLAP

La arquitectura ROLAP, accede a los datos almacenados en un *data warehouse* para proporcionar los análisis OLAP. La premisa de los sistemas ROLAP es que las capacidades OLAP se soportan mejor contra las bases de datos relacionales.

Utiliza una arquitectura de tres niveles. La base de datos relacional maneja los requerimientos de almacenamiento de datos, y el motor que utiliza proporciona la funcionalidad analítica. El nivel de base de datos usa bases de datos relacionales para el manejo, acceso y obtención del dato. El nivel de aplicación es el motor que ejecuta las consultas multidimensionales de los usuarios. El motor R-OLAP se integra con niveles de presentación, a través de los cuales los usuarios realizan los análisis OLAP. (9)

Proceso Analítico Multidimensional MOLAP

La arquitectura MOLAP usa bases de datos multidimensionales para proporcionar el análisis, su principal premisa es que el OLAP está mejor implantado almacenando los datos multidimensionalmente.

Un sistema de este tipo usa una base de datos propietaria multidimensional, en la que la información se almacena multidimensionalmente, para ser visualizada en varias dimensiones de análisis, a diferencia del ROLAP. Además utiliza una arquitectura de dos niveles: las bases de datos multidimensionales y el motor analítico. La base de datos multidimensional es la encargada del manejo, acceso y obtención del dato.

Las estructuras de almacenamiento son grandes arreglos dimensionales que son una copia de la fuente de datos y persisten físicamente en la misma estación de trabajo donde está instalada la herramienta *data warehousing*. Esto provoca que el acceso a la información almacenada se realice de forma más rápida y efectiva utilizándose en depósito donde el tiempo en la velocidad de respuesta es crítico. (9)

Proceso Analítico Híbrido HOLAP

El modo de almacenamiento HOLAP, como su nombre lo indica es un híbrido el cual combina las arquitecturas ROLAP y MOLAP para brindar una solución con las mejores características de ambas. Permite almacenar una parte de los datos como en un sistema MOLAP y el resto como en uno ROLAP.

Comparación entre los modos de almacenamiento ROLAP y MOLAP

Capítulo 1. Fundamentación teórica

Cada uno de los tipos de OLAP tiene beneficios en dependencia del problema en que se aplique. MOLAP requiere de menor espacio de almacenamiento y es más rápido calculando las agregaciones y devolviendo las respuestas, aunque se recomienda emplear para pequeños volúmenes de datos. ROLAP es considerado el más escalable, pero es más lento en el pre procesamiento y rendimiento de las consultas. HOLAP es rápido en el pre procesamiento y rendimiento de las consultas, aunque más lento que MOLAP y es escalable. (9)

ROLAP	MOLAP
Soportan análisis OLAP contra grandes volúmenes de datos elementales.	Se comportan razonablemente en volúmenes más reducidos (menos de 5 Gb).
Pueden crecer hasta un gran número de dimensiones.	Generalmente son adecuados para diez o menos dimensiones.
Las herramientas R-OLAP tienen menor rendimiento que las herramientas M-OLAP.	Las herramientas M-OLAP tradicionalmente tienen dificultades para consultar con modelos con dimensiones muy altas (del orden de millones de miembros).
Todos los datos de acceso están en la bodega almacenados.	Resúmenes de acceso a datos detallados en BD multidimensionales.
Vistas multidimensionales en la capa de presentación.	Tecnología propietaria para almacenar las vistas multidimensionales en arreglos no en tablas. Matriz de alta velocidad para la recuperación de los datos.
Las implementaciones ROLAP son más escalables y son frecuentemente atractivas a los clientes	En las implementaciones MOLAP el acceso a la información almacenada se realiza de forma

Capítulo 1. Fundamentación teórica

debido a que aprovechan las inversiones en tecnologías de bases de datos relacionales ya existentes en la organización.	más rápida y efectiva utilizándose un depósito donde el tiempo en la velocidad de respuesta es crítico. Normalmente se desempeñan mejor que la tecnología ROLAP, pero tienen problemas de escalabilidad.
---	--

Tabla 1. Comparación entre los modos de almacenamiento.

La selección de uno u otro modelo depende de cuán importante sea el rendimiento de las consultas para los usuarios y de la tecnología disponible a utilizar. En el modelo ROLAP la respuesta a las consultas y el tiempo de procesamiento suelen ser más lentos que con los modos de almacenamiento MOLAP o HOLAP. No obstante, ROLAP permite a los usuarios ver los datos en tiempo real y ahorrar espacio de almacenamiento al trabajar con conjuntos de datos grandes a los que no se suele consultar con frecuencia, como datos puramente históricos.

Para el desarrollo del mercado de datos se escogió como modo de almacenamiento de datos al Procesamiento Analítico Relacional (ROLAP) por las ventajas que el mismo significa en este tipo de entornos, donde el detalle y la consolidación de los distintos conceptos relacionados condicionan el funcionamiento del sistema. Además posee un ambiente conocido y disponibilidad de herramientas para su desarrollo lo cual facilita su puesta en práctica para el desarrollo de soluciones.

1.3.3. Almacenes de Datos (*Data warehouse*)

A partir de los años noventa surgen los almacenes de datos (*data warehouse*) debido a que las empresas necesitaban guardar grandes volúmenes de información, conservarla y consultarla en algún momento para poder llevar a cabo un análisis de la información almacenada, además de que los *data warehouse* sirven de base para la toma de decisiones.

Un Almacén de Datos es una base de datos corporativa que se caracteriza por integrar y depurar información de una o más fuentes distintas, para luego procesarla permitiendo su análisis desde infinidad de perspectivas y con grandes velocidades de respuesta. La creación de un *data warehouse* representa en la mayoría de las ocasiones el primer paso, desde el punto de vista técnico, para implantar una solución completa y fiable de *Business Intelligence*.

Capítulo 1. Fundamentación teórica

En el libro “*Mastering Data Warehouse Design Relational and Dimensional Techniques*” se precisa que los *data warehouse* producen una fuente estable de la información histórica que es constante, consistente y fiable para cualquier consumidor, además de que está configurado para facilitar los datos de cualquier forma de tecnología de análisis dentro de la comunidad de negocios. (10) La definición universalmente aceptada fue desarrollada por Bill Inmon, y dice que los *data warehouse* “Son un conjunto de datos orientados a un tema, integrados, de tiempo variante y no volátiles usados en la estrategia de toma de decisiones administrativas.” (11) Las características que tienen los *data warehouse* según Bill Inmon son las siguientes:

Integrado: Los datos almacenados en *data warehouse* deben integrarse en una estructura consistente.

Temático: Los datos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales.

Histórico: El almacén de datos se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones. Se utiliza para realizar análisis de tendencias.

No volátil: El almacén de datos puede ser leído pero no modificado. Es decir, la incorporación de los últimos valores que tomarán las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía.

Ventajas de un almacén de datos:

- Integrar datos históricos sobre la actividad de la organización (o negocio) en un único repositorio.
- Analizar los datos del negocio desde la perspectiva de su evolución en el tiempo.
- Prever tendencias de evolución del negocio.
- Identificar nuevas oportunidades de negocio y tomar decisiones estratégicas.
- Reducir los costes materiales y humanos en la toma de decisiones (12).

Desventajas del almacén de datos:

- Riesgo de fracaso en la construcción del sistema, al subestimar los costes de captura y preparación de los datos.
- Riesgo de fracaso en la construcción del sistema por cambios continuos en los requisitos de los usuarios.

Capítulo 1. Fundamentación teórica

- Problemas con la privacidad de los datos (12).

La implementación de un *data warehouse* permitirá a los funcionarios del MININT mejorar la disponibilidad de la información, disminuyendo el tiempo de obtención de la información, tendrán a la mano los resúmenes de información para armar reportes comparativos, utilizando los indicadores más comunes de la empresa, para apoyarse en la toma de decisiones; podrán quedar expuestos posibles defectos en aplicaciones actuales, siendo posible perfeccionarlas.

1.3.4. Modelo multidimensional

Para el diseño de un almacén de datos, varios autores han seguido diferentes conceptos para abordar el modelado de un almacén, entre los que se encuentran, esquema en estrella, esquema en copo de nieve, esquema en constelación entre otros. Independientemente de la metodología seguida, todos ellos utilizan un modelo multidimensional de datos.

El modelado multidimensional es: modelos de datos como conjuntos de medidas descritas por dimensiones.

- Adecuado para resumir y organizar datos (p.ej. hojas de cálculo).
- Enfocado para trabajar sobre datos de tipo numérico.
- Más fácil de visualizar y entender que el modelado E/R (13).

Características del modelo multidimensional

La estructura básica de un almacén de datos para el modelo multidimensional está definida por dos elementos: esquemas y tablas.

Como cualquier base de datos relacional, el almacén de datos se compone de tablas, de dos tipos de tablas específicamente en el modelo multidimensional.

- Tablas Facts o de Hechos: Tabla central en un esquema dimensional y en ella se guardan las medidas numéricas del negocio, por ejemplo: número de libros vendidos.
- Tablas Lock- up o dimensionales: Son las tablas que se alimentan de las tablas de hechos, por lo que contienen el detalle de los valores que se encuentran asociados a la tabla de hechos. (14)

La colección de tablas dentro del almacén se conoce como esquema y por lo general caen dentro de dos categorías básicas: esquema en estrellas y esquemas en copo de nieve.

Componentes del modelo multidimensional

Capítulo 1. Fundamentación teórica

El modelo multidimensional está compuesto por cubos, medidas, dimensiones, jerarquías, niveles, y atributos.

- Cubos: Son los principales objetos del proceso analítico en línea (OLAP, Online Analytic Processing), una tecnología que proporciona rápido acceso a los datos de un almacén de datos, conjunto de datos que normalmente se construyen a partir de un subconjunto de un almacén de datos y se organiza y resume en una estructura multidimensional definida por un conjunto de dimensiones y medidas.
- Dimensiones: Son un atributo estructural de los cubos y sirven como un mecanismo de selección de datos, organizadas en jerarquías de categorías y niveles que describen los datos de las tablas de hechos.
- Medidas: Dentro del modelo multidimensional, las medidas o atributos numéricos describen un cierto proceso del mundo real, el cual va a ser proceso de un análisis.
- Jerarquías: Es un conjunto de miembros de una dimensión y sus posiciones en relación con los demás, una forma de organizar los datos en diferentes niveles de agregación.
- Niveles: Es un elemento de la jerarquía de dimensiones. Describen la jerarquía desde el nivel de datos más altos (más resumido) hasta el más bajo (más detallado).
- Atributos: Provee información adicional acerca de los datos (15).

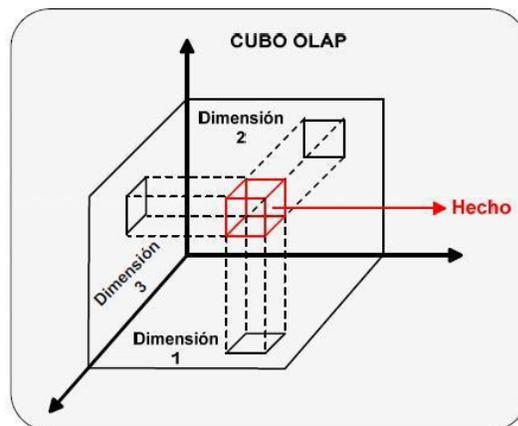


Figura 2. Estructura de un cubo OLAP.

Capítulo 1. Fundamentación teórica

1.3.5. Mercado de Datos (Datamart)

El *datamart* es una base de datos departamental o un subconjunto de una bodega de datos, especializada en el almacenamiento de los datos de un área de negocio específica, para un propósito específico. Se caracteriza por disponer la estructura óptima de datos para analizar la información al detalle desde todas las perspectivas que afecten a los procesos de dicho departamento. Un *datamart* puede ser alimentado desde los datos de un *data warehouse*, o integrar por sí mismo un compendio de distintas fuentes de información. Su función es apoyar a otros sistemas para la toma de decisiones. (16)

El motivo por el cual se crean mercados de datos es el crecimiento que tiene el almacén de datos y así facilitar su construcción y utilización. Las características de los mercados de datos son:

- Se centran en los requisitos de los usuarios asociados a un departamento o área de negocio concretos.
- Como diferencia con los almacenes de datos, los mercados no contienen datos operacionales detallados.
- Son más sencillos a la hora de utilizarlos y comprender sus datos, debido a que la cantidad de información que contienen es mucho menor que en los almacenes de datos. (16)

La creación de *datamart* permite agrupar los datos en áreas permitiendo simplificar el desarrollo del *data warehouse*, analizar la información al detalle desde varias perspectivas, logrando una mayor rapidez de consulta.

1.3.6. Integración de Datos

En aquellos escenarios donde exista un almacén de datos, este necesita ser alimentado de datos ya existentes en otras fuentes, y para ello aparecen los procesos de Extracción, Transformación y Carga (ETL por sus siglas en inglés).

- **Extracción:** Este es el primer paso para la obtención de la información de las diferentes fuentes hacia el almacén de datos. Como los datos provienen de fuentes distintas, generalmente se encuentran en formatos distintos y organizados a la forma de cada organización, la extracción deja los datos en un formato listo para la transformación.
- **Transformación:** Una vez que la información es extraída, el proceso de transformación se encarga de preparar los datos de la manera adecuada para integrarlos en el almacén de datos.

Capítulo 1. Fundamentación teórica

Para ello este proceso se compone de algunas actividades como: limpieza de datos, integración de formato, integración de datos.

- **Carga:** Una vez que los datos han sido extraídos de las diferentes fuentes y transformados, se procede a cargar el almacén de datos, después de la carga inicial se procede a mantener el almacén actualizándolo periódicamente (15).

Estos procesos son los que llevan el mayor peso en un almacén de datos permitiendo realizar la extracción de los datos de las fuentes de origen (base de datos del SUIN y los logs generados por el DGM), transformarlos y cargarlos hacia el *data warehouse*, limpiando los datos, desechando aquellos que no son de interés o poseen errores.

1.4. Metodología de desarrollo

Existen diferentes metodologías para el desarrollo de los almacenes de datos entre las que se destacan la metodología de Inmon y la metodología Kimball. Ralph Kimball y Bill Inmon son dos de las personalidades referentes y más influyentes en el área de los Almacenes de Datos. El primero es un especialista reconocido a nivel mundial en el diseño de los almacenes de datos y creador del enfoque Multidimensional; mientras que el segundo es el creador del término almacenes de datos y considerado como el padre de la disciplina.

La diferencia entre ambas metodologías viene dada por el enfoque que le da cada uno al problema.

La metodología de Inmon se basa en un enfoque *Top-Down* (Descendentemente), este enfoque propone construir primero el almacén de datos ya partir de este construir los mercados de datos. Propone la construcción de un repositorio de datos corporativo como fuente de información consistente, consolidada, histórica y de calidad. Como el almacén de datos se construye descendentemente los mercados de datos se nutren a partir del almacén corporativo, convirtiéndose en un complejo empresarial de bases de datos relacionales (17).

En comparación con la metodología que propone Inmon, la metodología de Kimball propone una arquitectura *Bottom-up* (Ascendentemente), plantea que se deben construir primeramente los mercados de datos orientados al tema de su departamento y luego el almacén de datos sería la unión de estos, lo que es eficaz y conduce a una solución completa en un corto período de tiempo.

En este caso, se decide utilizar la metodología Kimball por los siguientes elementos:

Capítulo 1. Fundamentación teórica

- Crea los conceptos de hechos y dimensiones, lo que indudablemente es muy eficaz en el proceso de la toma de decisiones y proporciona mayor agilidad en el proceso de desarrollo.
- Propone ir construyendo el almacén de datos a través de la construcción de los mercados de datos departamentales, lo que constituye una buena estrategia la que coincide con la división lógica de las empresas, entidades, organismos, etc.
- Es una metodología madura y reconocida por el resto de la comunidad dedicada al tema. Tiene bien definidas las etapas, actividades, artefactos y roles.
- Existe abundante documentación sobre la misma, la respuesta a todas las dudas y preguntas que puedan surgir se pueden encontrar en la web, a través de los servicios que brindan el grupo creador de la metodología.

Ciclo de vida que plantea la metodología



Figura 3. Ciclo de vida que plantea la metodología Kimball.

Planificación del Proyecto

La planificación busca identificar la definición y el alcance del proyecto de *data warehouse*, incluyendo justificaciones del negocio y evaluaciones de factibilidad. La planificación del proyecto se focaliza sobre

Capítulo 1. Fundamentación teórica

recursos, perfiles, tareas, duraciones y secuencialidad. El plan de proyecto resultante identifica todas las tareas asociadas con el ciclo de vida e identifica las partes involucradas.

Definición de los Requerimientos del Negocio

Los usuarios finales y sus requerimientos impactan siempre en las implementaciones realizadas de un *data warehouse*. Según la perspectiva de Kimball, los requerimientos del negocio se posicionan en el centro del “universo del *data warehouse*”. Como destaca siempre el autor, los requerimientos del negocio deben determinar el alcance del *data warehouse* (qué datos debe contener, cómo debe estar organizado, cada cuánto debe actualizarse, quiénes y desde dónde accederán, etc.).

Modelado Dimensional

Es realizado el modelo dimensional del *data warehouse* donde la definición de los requerimientos del negocio determina los datos necesarios para cumplir los requerimientos analíticos de los usuarios. Diseñar los modelos de datos para soportar estos análisis requiere un enfoque diferente al usado en los sistemas operacionales.

Diseño Físico

El diseño físico de las base de datos se focaliza sobre la selección de las estructuras necesarias para soportar el diseño lógico. Algunos de los elementos principales de este proceso son la definición de convenciones o estándares a seguir en la creación de la base de datos.

Diseño, Desarrollo y Procesamiento de Datos

Se definen como procesos de transformación los procesos para convertir o recodificar los datos fuente a fin de poder efectuar la carga efectiva del modelo físico. Por otra parte, los procesos de carga de datos son los procesos requeridos para poblar el *data warehouse*.

Diseño de la Arquitectura Técnica

Los ambientes de *data warehousing* requieren la integración de numerosas tecnologías. Se debe tener en cuenta tres factores: los requerimientos del negocio, los actuales ambientes técnicos y las directrices técnicas estratégicas futuras planificadas.

Selección de Productos e Instalación

Capítulo 1. Fundamentación teórica

Utilizando el diseño de arquitectura técnica como marco, es necesario evaluar y seleccionar componentes específicos de la arquitectura como por ejemplo: el motor de base de datos, la herramienta de ETL o el desarrollo pertinente, herramientas de acceso, etc.

Especificación de Aplicaciones para Usuarios Finales

No todos los usuarios del *warehouse* necesitan el mismo nivel de análisis. Es por ello que en esta etapa se identifican los roles o perfiles de usuarios para determinar los diferentes tipos de aplicaciones necesarias en base al alcance de los diferentes perfiles (gerencial, analista del negocio, vendedor, etc.)

Implementación

La implementación representa la convergencia de la tecnología, los datos y las aplicaciones de usuarios finales accesible desde el escritorio del usuario del negocio. Hay varios factores extras que aseguran el correcto funcionamiento de todas estas piezas, entre ellos se encuentran la capacitación, el soporte técnico, la comunicación, las estrategias de *feedback*¹.

Mantenimiento y crecimiento

Data warehousing es un proceso (de etapas bien definidas, con comienzo y fin, pero de naturaleza espiral) pues acompaña a la evolución de la organización durante toda su historia. Se necesita continuar con los relevamientos de forma constante para poder seguir la evolución de las metas por conseguir. Según afirma Kimball, “si se ha utilizado el ciclo de vida, el *data warehouse* está preparado para evolucionar y crecer”. Al contrario de los sistemas tradicionales, los cambios en el desarrollo deben ser vistos como signos de éxito y no de falla.

Gestión del Proyecto

La gestión del proyecto asegura que las actividades del ciclo de vida se lleven en forma y sincronizadas. Como lo indica el diagrama, la gestión acompaña todo el ciclo de vida. Entre sus actividades principales se encuentra el monitoreo del estado del proyecto y la comunicación entre los requerimientos del negocio y las restricciones de información para poder manejar correctamente las expectativas en ambos sentidos.

¹Retroalimentación a través de las acciones de los usuarios.

Capítulo 1. Fundamentación teórica

1.5. Herramienta de diseño y modelado

Embarcadero ER/Studio 8.0

Es una herramienta líder para el modelado de datos que permite documentar y reutilizar los activos en datos. Algunas de sus características son:

- Verdadera separación de modelos físicos y lógicos con integración completa, transformación automatizada y mapeo de tipos personalizados de datos.
- Generación de esquemas XML desde modelos físicos y lógicos.
- Soporte completo al ciclo de vida de bases de datos.
- Ingeniería inversa y directa.
- Generación automatizada de código de bases de datos.
- Almacén de datos y soporte a la integración.
- Validación de diseño, formación de la integridad referencial. (18)

Esta herramienta posee una interfaz clara y sencilla, permite realizar el modelado dimensional de los datos logrando así diseñar, documentar el modelo de datos para crear un entorno más estructurado y organizado, permite la realización cálculos de crecimiento de las tablas del *data warehouse* dando un indicador de los volúmenes de información que se manejan, además de ser la herramienta usada en el marco de trabajo del proyecto.

1.6. Herramientas para la realización de Inteligencia de Negocios

Oracle Business Intelligence Enterprise Edition

Oracle Business Intelligence Enterprise Edition (Oracle BIEE) es un conjunto completo de productos de inteligencia de negocios con los que provee un rango amplio de capacidades para el análisis y reporte de información. Se basa en una arquitectura moderna, altamente escalable y unificada que ofrece capacidades de inteligencia y análisis de datos a partir de distintas fuentes y aplicaciones.

Los componentes de OBIEE Plus son:

- Oracle BI Server: Servidor analítico
- Oracle BI Answers: Herramienta para consultas ad-hoc y análisis
- Oracle BI Interactive Dashboard: Cuadro de mando
- Oracle BI Publisher: Motor de informes.
- Oracle BI BriefingBooks: Capturador de informes

Capítulo 1. Fundamentación teórica

- Oracle BI Disconnected Analytics: Solución de movilidad que permite desconectarse de la red para trabajar con informaciones y cuadros de mando.
- Oracle BI Office Plug-In: Sincronización automática de Answers con Microsoft Word, Excel y PowerPoint.
- Oracle BI Delivers: Envío de alertas vía múltiples canales

Beneficios

- Funcionalidad de inteligencia de negocios completa, construida sobre una infraestructura unificada: Incluye tableros interactivos, consultas y análisis completos y a la medida (consultas no planeadas), alertas pre activas, reportes empresariales y financieros, análisis con procesamiento analítico en línea (OLAP), producción de reportes de alto volumen, inteligencia predictiva en tiempo real, análisis fuera de línea, así como integración con Microsoft Office.
- Conocimiento para todos y en donde sea: Provee conocimiento relevante para todos, no sólo para los analistas. Todos los niveles de la organización pueden ver información optimizada de acuerdo con el rol que desempeña cada persona.
- Modelo de negocio unificado: Construido sobre un modelo de información empresarial que unifica los metadatos a lo largo de las herramientas de BI de Oracle y aplicaciones analíticas con el costo total de pertenencia más bajo.
- Infraestructura de BI compatible: Se integra con cualquier fuente de datos, herramienta de extracción, transformación y carga, aplicación de negocios, servidor de aplicaciones, infraestructura de seguridad, tecnología de portales, u otros *front-ends*² y herramientas analíticas; bases de datos incluyendo IBM BD/2, Teradata, Microsoft SQL Server, SAP Business Information Warehouse, Microsoft Analysis Services, archivos planos, datos XML y datos no estructurados.

Esta herramienta permite la creación de reportes, tablas comparativas y análisis OLAP, ofreciendo a los funcionarios del MININT acceso al *data warehouse*, para que pueda hacer análisis avanzados, útiles para directivos y analistas, además de haber sido un requisito del cliente el uso de esta herramienta para la implementación de la aplicación.

²Término relacionado con el comienzo de un proceso.

Capítulo 1. Fundamentación teórica

1.7. Herramientas para la realización del proceso de ETL

1.7.1. Oracle Warehouse Builder (OWB)

OWB es una herramienta ETL producida por Oracle, que ofrece un entorno gráfico para construir, administrar y mantener la integración de datos en los procesos de inteligencia de negocios. Esta ya viene incluida en la instalación de las bases de datos Oracle a partir de la versión 11g.

Warehouse Builder ofrece las siguientes funcionalidades:

Almacenamiento de datos e Inteligencia de Negocios:

- Modelos relacionales y dimensionales.
- Soporte para OLAP y cubos organizados MVs (Materialized Views).
- Compatibilidad con características de almacenamiento de datos de Oracle como el particionamiento.
- Almacenamiento de datos específicos de los operadores para la carga de cubos y dimensiones.
- Soporte para el modelado y carga de SCD (*Slowly Changes Dimensions*³) tipo 2 y 3.
- Integración con herramientas como Oracle Discoverer y OBIEE (Oracle Business Intelligence Enterprise Edition).
- Integración de datos y ETL.
- Conectividad y opciones de movimiento de datos, incluyendo pasarelas, JDBC, y las cargas de datos a granel.
- Soporte para plataformas de bases de datos no nativas de Oracle.
- ELT con juego basado en Oracle SQL y basado en filas PL/SQL.
- Generación de código SQL nativo para bases de datos distintas de Oracle.
- Soporte para servicios web.

Esta herramienta se encuentra integrada con el sistema gestor de base de datos Oracle a partir de la versión 11g, por lo que se integra con dicho sistema gestor de base de datos permitiéndonos no solo la realización de las ETL, sino la implementación del *data warehouse*.

³Este concepto aplica a como una dimensión debe tener en cuenta los cambios históricos.

Capítulo 1. Fundamentación teórica

1.8. Sistema Gestor de Base de Datos

Oracle Database Enterprise Edition 11gR2

Es un sistema de gestión de base datos relacional desarrollado por Oracle Corporation. Se considera uno de los sistemas de bases de datos más completos, destacándose por su:

- Soporte de transacciones.
- Estabilidad.
- Escalabilidad.
- Soporte multiplataforma. (19)

Brinda una plataforma segura y escalable para el acceso confiable y rápido a todo tipo de información utilizando las interfaces estándar del sector. Permite una sólida administración del contenido para tipos avanzados de datos, esta versión incluye nuevas mejoras y características como son:

- La recopilación automática para lenguaje *PL/SQL*⁴ (Procedural Language / Structured Query Language) y Java en la base de datos.
- *Triggers* más rápidos, con inclusión de invocaciones más eficientes de *triggers* por fila.
- Operaciones SQL más rápidas y fáciles.
- Replicación *Oracle Data Guard*⁵ y *Oracle Streams*⁶ más rápida.
- Conexiones directas, más rápidas y confiables, a los dispositivos de almacenamiento del sistema de archivo de red.
- Actualizaciones más rápidas.
- Recuperación más rápida de archivos de gran tamaño. (20)

Lenguaje de programación PL/SQL

⁴Lenguaje de programación incrustado en Oracle.

⁵Oracle Data Guardes la funcionalidad de la base de datos Oracle que brinda la mayor y más efectiva disponibilidad, protección y recuperación ante desastres de los datos

⁶Oracle Streams propaga y administra datos, transacciones y eventos en una fuente de datos ya sea dentro de una base de datos, o de una base de datos a otra.

Capítulo 1. Fundamentación teórica

Este lenguaje es un gestor que va incorporado en Oracle para solventar las grandes limitaciones que existían con los editores SQL, en cuanto al control de las secuencias de ejecución de instrucciones, el uso de variables, la gestión de posibles errores.

Este lenguaje está basado en ADA⁷, por lo que incluye todas las características de los lenguajes de tercera generación. Esto permite manejar las variables, tener una estructura modular (procedimientos y funciones) y controlar las excepciones. Además incorpora un completo soporte para la programación orientada a objetos.

Los programas creados con PL/SQL pueden ser almacenados en base de datos como cualquier otro objeto, quedando disponibles para los usuarios. El uso del lenguaje PL/SQL es imprescindible para construir disparadores de bases de datos.

PL/SQL está incluido en el servidor siendo su lenguaje nativo y en algunas herramientas clientes. Soporta todos los comandos de consulta y manipulación de datos, aportando al lenguaje SQL las estructuras de control y otros elementos propios de los lenguajes de programación de tercera generación. (21)

1.9. Herramienta para la Gestión de la Base de Datos PLSQL Developer:

PL/SQL Developer es un ambiente integrado para el desarrollo, prueba, depuración de errores y optimización de PL/SQL de Oracle, almacenado en unidades de programa como paquetes y *triggers*, entre otros. PL/SQL Developer contiene ayuda sensitiva al contexto, descripciones de bases de datos de objetos, sintaxis resaltada, edición y búsqueda de datos, browser gráfico y muchas otras características que le hacen la vida más fácil al usuario. (22)

1.10. Conclusiones

Lo planteado en este capítulo sirve de base para desarrollar la solución de inteligencia de negocios para el Sistema Único de Identificación Nacional, para ello se abordaron conceptos asociados al tema, se realizó una caracterización de la metodología, herramientas y tecnologías a utilizar en el desarrollo de la solución donde se arribaron a las siguientes conclusiones.

⁷Ada es un lenguaje de programación orientado a objetos, fuertemente tipado y concurrente.

Capítulo 1. Fundamentación teórica

El uso de la metodología Kimball permitirá agilizar el proceso de desarrollo permitiendo brindar soluciones parciales al cliente. Se fundamentan las tecnologías definidas por el equipo de trabajo en coordinación con los clientes atendiendo a las ventajas que ofrecen estas al desarrollo integrado.

Capítulo 2. Análisis y diseño de la solución

Capítulo 2. Análisis y diseño de la solución

2.1. Introducción

En este capítulo se abordarán un grupo de elementos que se deben tener en cuenta en el desarrollo de la solución y para un mejor entendimiento del negocio. Se definirán los temas de análisis, roles, permisos y necesidades del usuario. También se definirán los requisitos de información, requisitos funcionales y no funcionales. Se confeccionará además el modelo de datos y la matriz bus, se identificarán los hechos y dimensiones del *data warehouse*.

2.2. Descripción del negocio

El Sistema Único de Identificación Nacional de la República de Cuba es una solución desarrollada con un enfoque a procesos y concebida como el eje fundamental que integra tecnologías para la verificación biométrica, aspecto clave para la creación de un sistema de identificación moderno.

Está diseñada bajo el precepto de la creación de una Base de Datos Nacional, poblada a partir de la información vigente y de la información resultante de los distintos trámites que se realizan en las oficinas de identificación y registros, con un enfoque de mejora y depuración continua, que garantizará el manejo de una identidad única para cada persona natural o extranjera residente en el país.

Este sistema permite la gestión de los diferentes procesos requeridos para la ejecución segura y eficiente de trámites de identificación, incluyendo la búsqueda en bases de datos, la captura de los datos alfanuméricos, documentales y biométricos, la validación de la identidad por comparación biométrica, y la obtención de reportes para este nivel.

Permite el registro de las trazas de las acciones realizadas, y se responsabiliza además de la interacción con AFIS para la validación a partir de rasgos biométricos, así como de la comunicación con sistemas o servicios externos que se requieran para el correcto funcionamiento de los procesos.

La obtención de reportes se realiza a través del propio sistema, consultando directamente la base de datos nacional, donde la información se encuentra disgregada en más de 255 tablas, donde algunas de las mismas poseen millones de *records*⁸ que son producto de los más de 6000 trámites diarios que se realizan en las oficinas de identificación y registro, lo cual trae consigo una demora de 2 – 10 minutos en para la obtención de los mismos.

⁸Filas que existen en una tabla.

Capítulo 2. Análisis y diseño de la solución

La Dirección de Identificación y Registros genera una serie de reportes no solo para análisis del MININT, sino también para instituciones como las FAR (Fuerzas Armadas Revolucionarias), el Banco Nacional de Cuba y la ONE (Oficina Nacional de Estadísticas).

2.3. Necesidades del usuario

Las necesidades de los usuarios son de vital importancia para la creación de productos y servicios de éxito. Las necesidades de los clientes son diversas y es importante que se entienda que si se desea mejorar o crear valor, los usuarios evalúan el desempeño de los productos y servicios a través de sus características.

Los reportes que se brinden deben cumplir con la siguiente pirámide informativa:



Figura 4. Pirámide informativa.

Los reportes deben dar una visión macro de la información permitiendo la navegabilidad de los datos o sea brindando la posibilidad de llegar desde el nivel más alto al bajo.

Se deben brindar reportes o salidas informativas:

- Salidas con estadísticas sobre la información de personas que se brinda a la FAR, ONE, Banco y Trabajo y Seguridad Social u otros organismos.
- Estadísticas de población:
 - Por provincia, municipio, consejo popular y circunscripción.
 - Por sexo.
 - Por rango de edad.
 - Por color de la piel.

Capítulo 2. Análisis y diseño de la solución

- Movimientos migratorios, dentro del propio municipio, entre municipios de la propia provincia y entre municipios de diferentes provincias. General para residentes permanentes o transitorios y de ellos con antecedentes que se controlan por los órganos del MININT.
- Salida de información de personas con antecedentes policiales que realizan cambio de domicilio a la PNR (Policía Nacional Revolucionaria), estadísticas municipio y provincia, que incluye las zonas especiales que se definan.
- Desempeño laboral:
 - Por unidades.
 - Por funcionarios.

2.4. Reglas de Negocio

En la fase de análisis una de las tareas más importante son las reglas de negocio, las mismas describen políticas que deben cumplirse o condiciones que deben satisfacerse, por lo que regulan algún aspecto del negocio. También dan la oportunidad de definir cómo es que se van a calcular las medidas del almacén de datos.

A continuación se definen un conjunto de indicadores⁹ identificados y definidos por un grupo de ingenieros industriales de la DIR:

Indicador	Fórmula de cálculo
Nivel de fiabilidad del servicio	$(1 - \text{cantidad de irregularidades detectadas}) / \text{total de trámites realizados} * 100$
Coficiente de fiabilidad del servicio	$\text{Cantidad de trámites con irregularidades detectadas} / \text{Total de solicitudes de servicio}$
Tiempo promedio de servicio en la Unidad CIRP	Tiempo promedio de los trámites realizados.
Nivel de estabilidad en el tiempo de servicio en la Unidad CIRP	$(\text{Cantidad de solicitudes de servicio que son atendidas en el tiempo establecido}) / \text{Total de}$

⁹Se refiere a los datos esencialmente cuantitativos referentes a los objetivos y tareas que propone una organización.

Capítulo 2. Análisis y diseño de la solución

	solicitudes de servicio) * 100
Tiempo promedio de respuesta a las irregularidades	\sum Tiempo en que se le da respuesta a la irregularidad / Total de irregularidades atendidas.
Nivel de estabilidad en el tiempo de respuesta a las irregularidades	(Cantidad de irregularidades a las que se les da respuesta en el tiempo establecido / Total de irregularidades atendidas) * 100

Estos indicadores se encuentran descritos en el documento Sistema de indicadores organizacionales del CMI.

A continuación se mostrarán las reglas de negocio identificadas para la realización del *data warehouse* del Sistema Único de Identificación Digital:

- Los identificadores de las tablas no pueden tomar valores repetidos.
- Las medidas de los hechos deben poseer valores mayores o iguales a cero, nunca un número negativo.
- Además forman parte de estas reglas de negocio cada uno de los indicadores descritos.

2.5. Especificación de requisitos

Después de realizar un análisis de los principales indicadores y requisitos informativos para el Sistema Único de Identificación Nacional se identificaron los requisitos necesarios para satisfacer las necesidades de los clientes.

2.5.1. Requisitos de información

Los requisitos de información (RI) describen qué información debe almacenar el sistema para satisfacer las necesidades de clientes y usuarios. Identifican los conceptos relevantes sobre los que se debe almacenar información y los datos específicos que son de interés. A continuación se muestran los requisitos de información identificados durante el proceso de análisis:

RI 1. Obtener trámites realizados por cantidad y por tipos.

RI 2. Obtener nivel de fiabilidad del servicio.

Capítulo 2. Análisis y diseño de la solución

- RI 3.** Obtener coeficiente de fiabilidad del servicio.
- RI 4.** Obtener tiempo promedio de servicio en la Unidad del Carnet de Identidad y Registro de la Población (CIRP).
- RI 5.** Obtener nivel de estabilidad en el tiempo de servicio en la Unidad del CIRP.
- RI 6.** Obtener tiempo promedio de respuesta a las irregularidades.
- RI 7.** Obtener nivel de estabilidad en el tiempo de respuesta a las irregularidades.
- RI 8.** Obtener reportes de la población por:
 - RI 8.1.** Por provincia, municipio, consejo popular y circunscripción.
 - RI 8.2.** Por sexo.
 - RI 8.3.** Por rango de edad.
 - RI 8.4.** Por color de la piel.

2.5.2. Requisitos funcionales

Los requerimientos funcionales de un sistema describen la funcionalidad o los servicios que se espera que éste provea. Son entendidos como capacidades que debe exhibir una aplicación con el fin de resolver un problema. Entre los identificados se encuentran:

- RF 1.** Extraer, transformar y cargar datos de los sistemas fuentes.
- RF 2.** Mostrar reporte.
- RF 3.** Adicionar reporte.
- RF 4.** Eliminar reporte.
- RF 5.** Modificar reporte.
- RF 6.** Exportar reporte a PDF.
- RF 7.** Exportar reporte a Excel.

Ver descripción de los requisitos funcionales en el [Anexo 1](#).

2.5.3. Requisitos no funcionales

Los requisitos no funcionales describen aquellas características no funcionales que los clientes y usuarios desean que tenga el sistema a desarrollar. Entre los identificados se encuentran:

Software

Capítulo 2. Análisis y diseño de la solución

- RNF 1.** Usar Linux Suse 10 SP 2 o superior.
- RNF 2.** Usar servidor de Base de Datos Oracle 11gR2.
- RNF 3.** Usar servidor web Oracle WebLogic 10.3.2.
- RNF 4.** Las PC clientes deben tener instalado un navegador web que posea cualquiera de los siguientes motor/intérprete de HTML, CSS y JavaScript:
- Trident (Internet Explorer a partir de la versión 7).
 - WebKit (Konkeror, Safari, Google Chrome).
 - Gecko (Mozilla Firefox, Flock).
 - Presto (Opera).
- RNF 5.** Servidor de base de datos con las siguientes características:
- 4 GB de RAM o superior.
 - 160 GB de capacidad en disco duro o superior.
 - Procesador 2GHz o superior.
- RNF 6.** Servidor web con las siguientes características:
- 2 GB de RAM o superior.
 - 40 GB de capacidad en disco duro o superior.
 - Procesador 2GHz o superior.

Diseño e implementación.

- RNF 7.** Usar Embarcadero ER\Studio 8.0 para el modelado.
- RNF 8.** Usar PL/SQL Developer para la gestión de la base de datos.
- RNF 9.** Usar Oracle Business Intelligence 10.1 para gestionar los reportes.
- RNF 10.** Usar PL/SQL como lenguaje de consultas.

Interfaz de usuario

Los requerimientos de interfaz de usuario se centran en la distribución de contenido y presencia de la información de cara al cliente.

- RNF 11.** Los reportes deben contar con una interfaz simple que facilite la interacción del usuario con la aplicación.
- RNF 12.** Las interfaces de salida no serán cargadas con información innecesaria.

Capítulo 2. Análisis y diseño de la solución

RNF 13. Permitir la navegabilidad entre los datos, logrando navegar entre los distintos niveles de información.

Interfaz hardware

RNF 14. Interfaz capacitada para conectarse a los dispositivos de impresión.

Usabilidad

Por las características de los nuevos sistemas de almacenamiento y análisis de los datos, en específico los almacenes de datos, los usuarios finales deben tener conocimiento de las herramientas de integración de los datos. Se debe realizar una capacitación al personal sobre el uso de los almacenes de datos, para la realización de los reportes y el entendimiento de los valores reflejados en ellos. Para lograr esto se debe lograr que el sistema tenga las funcionalidades adecuadas.

RNF 15. El sistema debe tener una interfaz gráfica uniforme.

RNF 16. Se necesita que los usuarios normales y avanzados sean productivos operando el sistema en poco tiempo hasta que la organización logre un estado estable para todos.

RNF 17. El sistema debe ser fácil de usar para los usuarios.

RNF 18. El usuario podrá acceder de manera rápida a la información que solicita en el área pertinente de acuerdo al objetivo de la solicitud.

RNF 19. El sistema será distribuido en idioma español, aunque estará preparado para soporte multilinguaje.

Seguridad

RNF 20. El sistema debe permanecer disponible las 24 horas del día. En caso de fallo, pudiera estar fuera de servicio por un período máximo de 48 horas.

RNF 21. El tiempo de mantenimiento del sistema lo definirá el administrador del sistema y en un horario que no afecte la productividad, preferentemente en horarios nocturnos.

RNF 22. El sistema registrará todas las acciones que se realizan.

RNF 23. La conexión a los servidores de bases de datos y de aplicaciones es requerida para el correcto funcionamiento.

Estándares aplicables

Capítulo 2. Análisis y diseño de la solución

RNF 24. En cada etapa de desarrollo del sistema, el resultado parcial es sometido a pruebas de liberación.

RNF 25. Para el diseño de la base de datos se utilizará un convenio definido para los nombres de las tablas, vistas y funciones que sean necesarias implementar.

2.6. Modelo de arquitectura a seguir

Componentes del *data warehouse*: Los almacenes de datos están integrados por una serie de procesos que en su conjunto definen el ambiente que estos poseen.

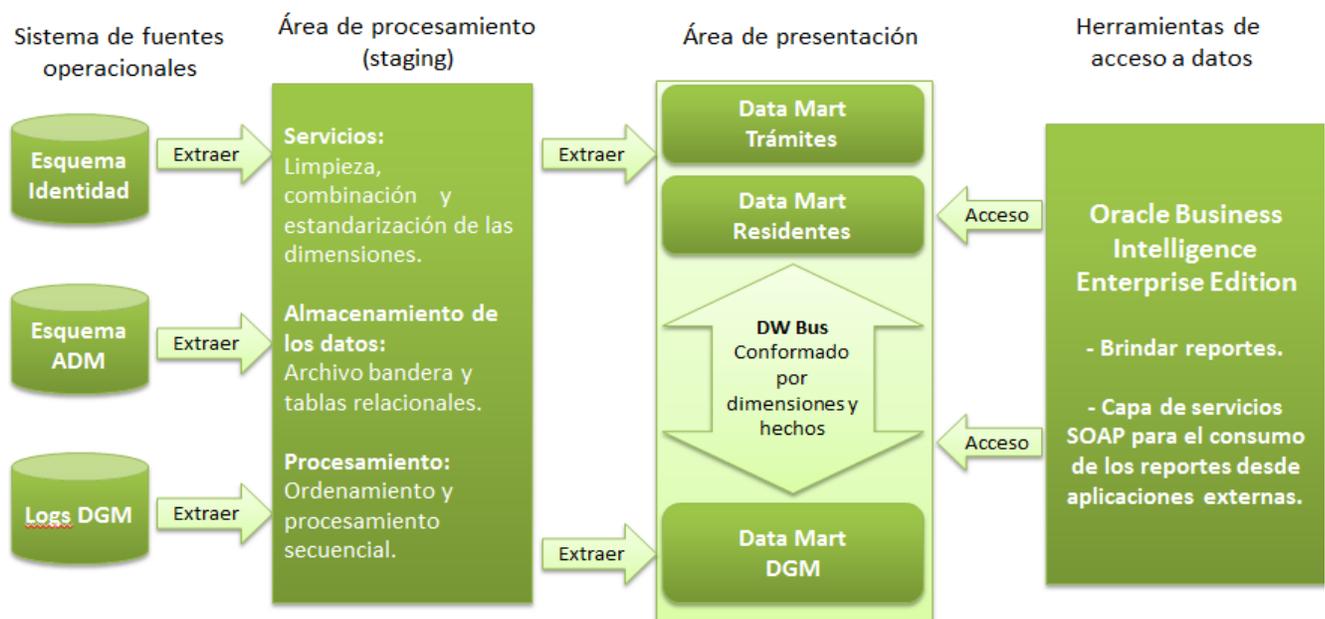


Figura 5. Componentes que conforman el *data warehouse*.

Sistema de fuentes operacionales

Se definen las fuentes de donde serán obtenidos los datos. En este caso los sistemas fuentes son:

- Base de datos de identificación nacional (esquema identidad). (23)
- Base de datos de identificación nacional (esquema *ADM*¹⁰). (24)

¹⁰Esquema de base de datos correspondiente al módulo de administración del SUIN.

Capítulo 2. Análisis y diseño de la solución

- Logs generados por el DGM¹¹. Ver modelo de dominio en el [Anexo 2](#).

Área de procesamiento (*staging*)

Es el área que se encarga del almacenamiento de los datos temporalmente y de realizar un conjunto de procesos comúnmente llamados de Extracción, Transformación y Carga (ETL, por sus siglas en inglés). Realiza la función de interfaz entre las fuentes operacionales y el área de presentación. En este componente es donde se invierte la mayor cantidad de tiempo y esfuerzo durante el desarrollo del almacén.

Se realiza el proceso de extracción de los datos de los esquemas ADM e Identidad y de los logs del DGM, almacenándose temporalmente en la base de datos para su posterior análisis y procesamiento. Luego de ser almacenados se procede a la detección de inconsistencias, duplicaciones, errores de formato e inexistencias, estandarizándose la información almacenada en las diferentes fuentes. Estas transformaciones son las que sirven de sostén para realizar la carga de los datos, hacia el almacén de datos, en el área de presentación.

A continuación se muestra un diseño de cómo se realizará el proceso explicado anteriormente:

¹¹Device Grid Manager, software encargado de la gestión centralizada de dispositivos.

Capítulo 2. Análisis y diseño de la solución

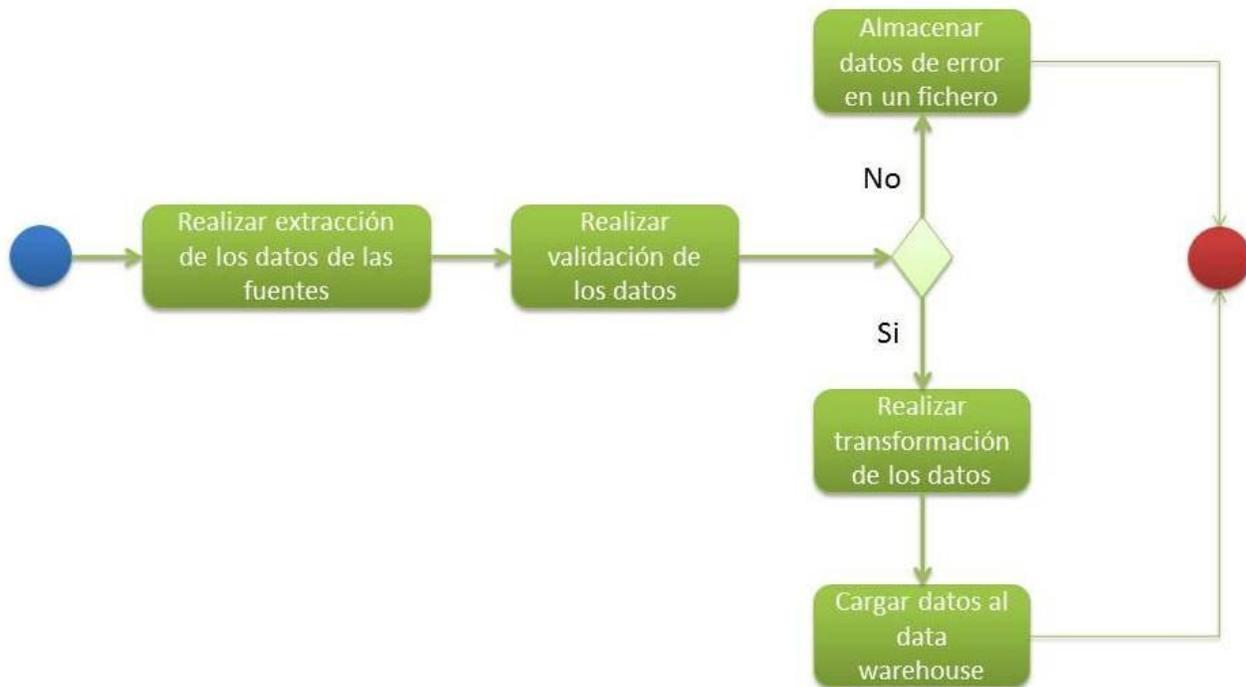


Figura 6. Diagrama de flujo del proceso de integración de los datos.

Área de Presentación

Es el componente donde los datos se encuentran organizados, almacenados y disponibles para ser consultados, reportados o analizados por parte de los usuarios finales. La información, diseñada mediante esquemas dimensionales, que ha sido definida por los usuarios como útil para la toma de decisiones. Generalmente esta área es referenciada como una serie de *datamart* integrados donde cada uno se encuentra representando a un proceso específico del negocio.

Se han definido tres *datamart*:

- Trámites: Se guardan los datos correspondientes a los trámites.
- Residentes: Se guardan los datos correspondientes las personas residentes en el país.
- DGM: Se guardan los datos correspondientes al manejo de los dispositivos de captación de datos biométricos.

Capítulo 2. Análisis y diseño de la solución

Herramientas de Acceso a Datos: En este componente es utilizada la palabra herramientas para referirse a la variedad de capacidades que pueden ser suministradas a los usuarios del negocio para dar soporte a la toma de decisiones, teniendo como actividad principal la posibilidad de consultar el área de presentación del almacén de datos.

Como herramienta de acceso a datos fue definido Oracle Business Intelligence Enterprise Edition permitiendo así la obtención de reportes en tiempo real, además de una capa de servicios SOAP¹² que permite obtener los reportes definidos para su uso desde aplicaciones externas.

2.7. Modelo de datos Dimensional

Dentro de la modelación de datos se encuentra el Modelo Dimensional, el cual está compuesto por hechos y dimensiones.

Los hechos son los datos que brindan información cuantitativa sobre las características del negocio que se quieren analizar, en este caso sobre los temas de análisis identificados para cada área. Por otra parte, las dimensiones serían los grupos de elementos que permitirían analizar esos valores contenidos en los hechos.

Este tipo de modelado posee la técnica de diseño para estructurar los datos según el tema de clasificación, se basa en la simplicidad, es decir, es más intuitivo para el usuario y está optimizado para lograr un mejor rendimiento de las consultas.

2.7.1. Matriz Bus

La Matriz Bus refleja gráficamente la relación que existe entre las dimensiones y hechos del *data warehouse*, a continuación se presenta la tabla de dicha relación, clasificada por los temas de análisis definidos.

A continuación se listan las dimensiones, hechos y medidas identificadas:

Dimensiones:

- DimTramite: Se describe el trámite que se realizó.
- DimPersona: Se registran cada una de las personas del país. Posee la siguiente jerarquía:
 - Provincia → Municipio → Persona

¹²Es un protocolo estándar que define cómo dos objetos en diferentes procesos pueden comunicarse por medio de intercambio de datos XML. (29)

Capítulo 2. Análisis y diseño de la solución

- DimOficina: Describe la oficina que realiza el trámite. Posee la siguiente jerarquía:
 - Provincia → Oficina
- DimTiempo: Se usa para ubicar los datos en un periodo de tiempo determinado. Posee la siguiente jerarquía:
 - Año → Trimestre → Mes → Día
- DimTipoDir: Representa el tipo de dirección en cada uno de los cambios de dirección.
- DimDirActual: Representa la dirección para la cual la persona cambio. Posee la siguiente jerarquía:
 - Provincia → Municipio → Dirección
- DimDirProcedencia: Representa la dirección de la cual la persona proviene. Posee la siguiente jerarquía:
 - Provincia → Municipio → Dirección
- DimDispositivo: Describe el dispositivo que genera los logs.
- DimEstacion: Describe la estación de trabajo donde fueron generados los logs.
- DimUbicacion: Representa la ubicación. Posee la siguiente jerarquía:
 - Provincia → Municipio

Hechos:

- FactTramite: Encargado de registrar los trámites realizados, vincula las dimensiones: DimTramite, DimOficina y DimTiempo, posee las medidas: cantIrregularidades y tiempoTramite.
- FactPersona: Encargado de llevar los indicadores referentes a las personas, vincula las dimensiones: DimPersona y DimUbicacion, posee las medidas: cantidad, menorEdad, mayorEdad, cpblanca, cpnegra, cpmestiza, varon, hembra.
- FactCambioDir: Encargada de llevar el control de los cambios de dirección, vincula las dimensiones: DimFuncionario, DimOficina, DimIrregularidad y DimTiempo, posee la medida: cantidad.
- FactTramitesOficina: Posee los indicadores referentes a los trámites realizados por las oficinas, vincula las dimensiones: DimOficina, DimTiempo, posee las medidas: cantIrrDetec, cantTra, tiempoProServ, nivelEstabTiempo, tiempoPormResIrr, nivelEstabTiempoIrr.
- FactDispositivo: Encargado de llevar el tiempo de uso de los dispositivos, vincula las dimensiones: DimEstacion, DimDispositivo, DimTiempo, DimOficina, posee la medida: tiempoUso.

Capítulo 2. Análisis y diseño de la solución

Medidas:

- cantIrregularidades: Recoge la cantidad de irregularidades encontradas en un trámite.
- tiempoTramite: Guarda el tiempo en el que se realiza el trámite.
- tiempoEstado: Almacena el tiempo que tarda el trámite en ese estado.
- cantPersona: Recoge la cantidad de personas.
- tiempoRespuesta: Guarda el tiempo de respuesta a una irregularidad.

Dimensiones\Hechos	FactCambioDir	FactDispositivo	FactPersona	FactTramite	FactTramites Oficina
DimDirActual	X				
DimDirProcedencia	X				
DimDispositivo		X			
DimEstacion		X			
DimOficina	X	X		X	X
DimPersona	X		X		
DimTiempo	X	X		X	X
DimTipoDir	X				
DimTramite				X	
DimUbicacion			X		

Tabla 2. Matriz Bus

2.7.2. Modelo de datos

El modelo de datos está compuesto por los hechos, dimensiones y medidas definidas. Para una información más detallada ver artefacto Modelo de Datos y la vista ampliada en el [Anexo 3](#).

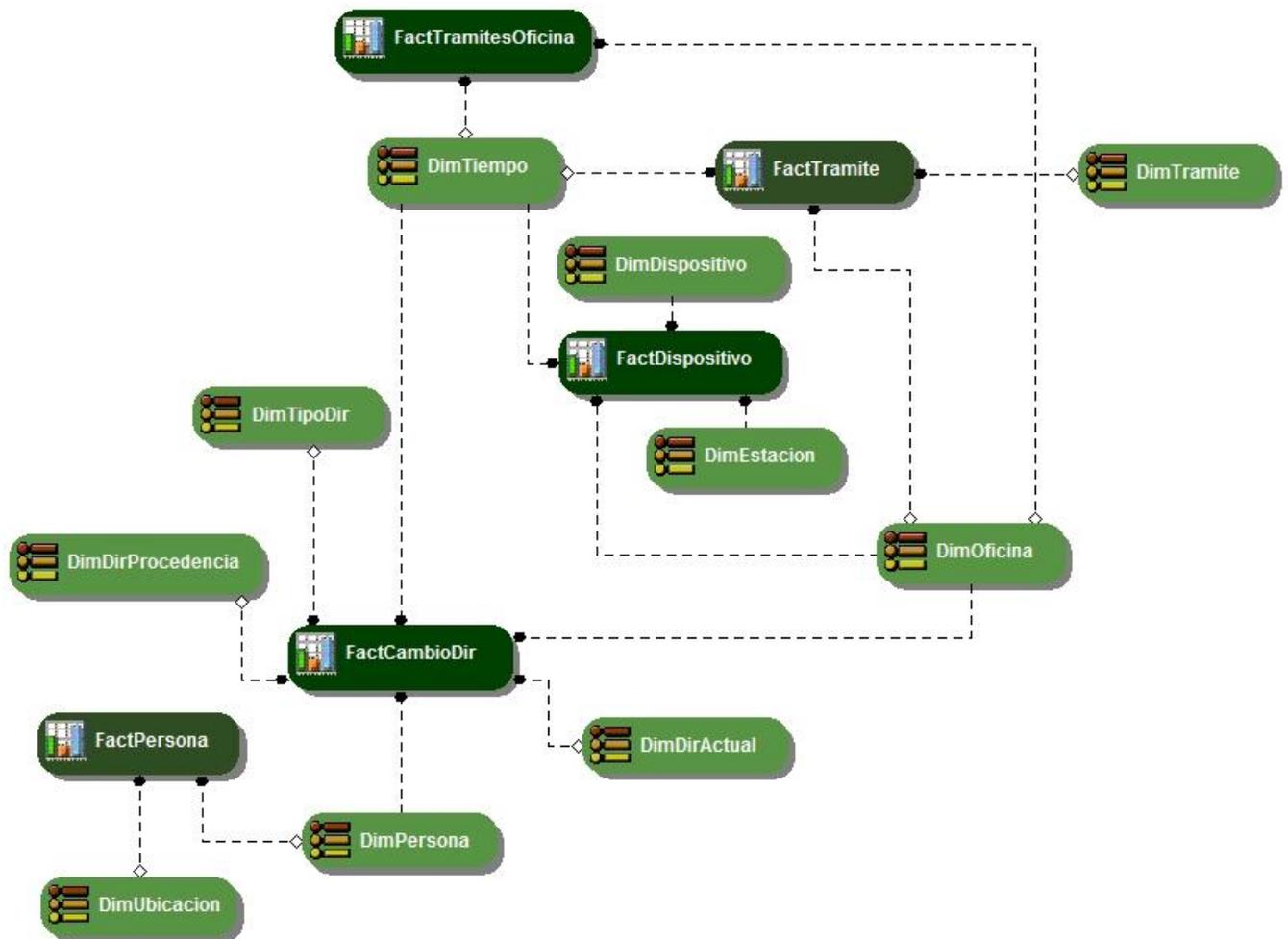


Figura 7. Modelo de datos.

2.8. Esquema de Seguridad

Representa el respaldo por los niveles de acceso, específicamente por los roles definidos. Administra los permisos que tendría cada tipo de usuario a la hora de acceder a la información del sistema.

A continuación se muestran los roles y permisos que los usuarios poseen en su interacción con la base de datos y la aplicación.

Seguridad en base de datos

Rol definido para la interacción de los usuarios con la base de datos:

Capítulo 2. Análisis y diseño de la solución

Rol	Descripción
Administrador de ETL	Realiza el proceso de ETL de los datos.
Administrador	Tiene permisos de insertar, modificar y eliminar.
Consulta	Solo posee permiso de consulta.

Tabla 3. Seguridad en la base de datos.

Seguridad en la aplicación

A continuación se definen los roles de seguridad para la interacción de los usuarios con la solución de inteligencia de negocio para hacer sostenible el manejo de la aplicación.

Rol	Descripción
Administrador	Posee permisos para realizar cualquier tipo de acción en la aplicación.
Administrador de Reportes	Encargado de la realización de los reportes.
Analista de información	Visualizará los reportes.

Tabla 4. Seguridad en la aplicación.

Dentro de los analistas de información se encuentran:

- Jefes de la Dirección de Identificación y Registros.
- Jefes del Departamento CIRP.
- Jefes Provinciales SIR.

Conjuntamente con los niveles de accesos y roles antes determinados, para interactuar con la aplicación, la plataforma Oracle Business Intelligence tiene incluida su propia seguridad, la cual se divide en cuatro áreas fundamentales, estas son:

Seguridad de acceso a datos de objetos: Incluye usuarios, contraseñas, autorizaciones permitidas, recursos web y protección a datos.

Autenticación: Tiene que ver con el procesamiento de información interactiva de inicio de sesión (por ejemplo nombre de usuario y contraseña) comparándola con la información recuperada del almacén de datos de seguridad.

Capítulo 2. Análisis y diseño de la solución

Autorización de recursos web (URL): Brinda protección a las URL para responder a cada usuario si pueden o no acceder a una determinada página. Esto es decidido por el administrador de recursos web, el cual le brinda a cada usuario autenticado un permiso de seguridad, delimitando las páginas a las que tiene acceso y a las que no.

Autorización a objetos del dominio: En el sistema los únicos objetos del dominio protegidos por la plataforma son los objetos de repositorio otorgados al usuario autenticado. Es responsabilidad de los objetos del dominio autorizar las operaciones solicitadas por este.

2.9. Conclusiones

En el presente capítulo se realizó un estudio del negocio y se analizaron las necesidades de los usuarios. Se realizó un levantamiento de requisitos donde se tomaron 8 requisitos de información, 7 requisitos funcionales y 24 no funcionales. Se planteó la arquitectura a seguir para la construcción del *data warehouse*, fueron definidos los hechos y dimensiones que conforman la matriz bus, quedando definido el modelo de datos para el almacén de datos. Se conformó el esquema de seguridad para la aplicación donde se especificaron los roles de los usuarios. Se sentaron las bases que constituyen la entrada para la implementación de la solución.

Capítulo 3. Implementación de la solución

Capítulo 3. Implementación de la solución

3.1. Introducción

En este capítulo se realiza la implementación de la solución de inteligencia de negocios para el Sistema Único de Identificación Nacional, realizándose el proceso de ETL y almacenamiento de los datos en las tablas correspondientes de acuerdo al modelo de datos definido en el capítulo anterior, posteriormente se define la capa de inteligencia de negocios y con ella la implementación de los reportes candidatos que responden a las necesidades del cliente, permitiendo visualizar los datos a través de textos, tablas y gráficos que permiten al usuario un mejor entendimiento y comprensión de los mismos.

3.2. Implementación de la base de datos

Tablas

El *data warehouse* del SUIN cuenta con 15 tablas, 10 dimensiones y 5 hechos, que a continuación se muestran:

	Tablas
Dimensiones	DimDirActual
	DimDirProcedencia
	DimDispositivo
	DimTramite
	DimUbicacion
	DimEstacion
	DimOficina
	DimPersona
	DimTiempo
	DimTipoDir
Hechos	FactTramite
	FactPersona
	FactDispositivo
	FactCambioDir
	FactTramitesOficina

Tabla 5. Hechos y Dimensiones.

Secuencias

Son los atributos que se van a ir incrementando secuencialmente durante la entrada de los datos a la base de datos, en este caso las llave primarias. Las secuencias identificadas son las siguientes:

Capítulo 3. Implementación de la solución

DIMENSION_KEY_PK_1	DimTipoDir	DIMENSION_KEY
DIMTRAMITE_DIMENSION_KEY_PK	DimTramite	DIMENSION_KEY
NTRAMITE_IDX	DimTramite	IDPROCESO
DIMENSION_KEY_PK	DimUbicacion	DIMENSION_KEY
NMUNICIPIO_IDX	DimUbicacion	IDMUNICIPIO
NPROVINCIA_IDX_1	DimUbicacion	IDPROVINCIA
DIMDIRACTUAL_IDX	FactCambioDir	DIMDIRACTUAL
DIMDIRPROCEDENCIA_IDX	FactCambioDir	DIMDIRPROCEDENCIA
DIMOFICINA_IDX_2	FactCambioDir	DIMOFICINA
DIMPERSONA_IDX_1	FactCambioDir	DIMPERSONA
DIMTIEMPO_IDX_2	FactCambioDir	DIMTIEMPO
DIMTIPODIR_IDX	FactCambioDir	DIMTIPODIR
DIMDISPOSITIVO_IDX	FactDispositivo	DIMDISPOSITIVO
DIMESTACION_IDX	FactDispositivo	DIMESTACION
DIMOFICINA_IDX_3	FactDispositivo	DIMOFICINA
DIMTIEMPO_IDX_3	FactDispositivo	DIMTIEMPO
DIMPERSONA_IDX	FactPersona	DIMPERSONA
DIMUBICACION_IDX	FactPersona	DIMUBICACION
DIMOFICINA_IDX_1	FactTramitesOficina	DIMOFICINA
DIMTIEMPO_IDX_1	FactTramitesOficina	DIMTIEMPO
DIMUBICACION_IDX_2	FactTramitesOficina	DIMUBICACION
DIMOFICINA_IDX	FactTramite	DIMOFICINA
DIMTIEMPO_IDX	FactTramite	DIMTIEMPO
DIMTRAMITE_IDX	FactTramite	DIMTRAMITE
DIMUBICACION_IDX_1	FactTramite	DIMUBICACION

Tabla 7. Índices.

3.2.1. Implementación de los procesos de ETL

El proceso de ETL es el más importante, pues es quien organiza el flujo de los datos y aporta los métodos y herramientas necesarios para mover datos desde múltiples fuentes a un almacén de datos, reformatearlos, limpiarlos y cargarlos en otra base de datos o mercado de datos.

La primera parte del proceso consiste en extraer los datos desde el sistema origen, una parte importante de este proceso es analizarlos, de lo que resulta un chequeo que verifica si estos cumplen la pauta o estructura que se esperaba, de no ser así, son rechazados. En el caso de los datos procedentes de la

Capítulo 3. Implementación de la solución

base de datos fuente ya se encuentran listos para pasar al proceso de transformación. En esta etapa del proceso no fue necesario realizar una limpieza de los datos, pues los mismos no poseen inconsistencias.

El proceso de transformación se realizó de acuerdo a las reglas del negocio que se definieron, para su realización se ejecutaron 19 transformaciones. Finalmente está el proceso de la carga, este es el momento cuando los datos de la fase anterior son cargados al sistema destino.

Para los procesos de ETL nos apoyamos en el flujo descrito en el capítulo anterior ([Figura 5](#)).

Los datos son obtenidos a partir de la base de datos Identidad y a partir de los *logs* generados por el DGM (*Device Grid Manager*) los cuales son almacenados en base de datos. Para asegurar que no sean cargados todos los datos en cada ejecución de la ETL; primeramente se realiza una carga de todos los datos, posteriormente solo son extraídos los datos que fueron insertados o modificados a partir de la última fecha que fue realizada la carga de los datos, para garantizar lo planteado es guardada la fecha en que se realiza la carga de los datos de las personas, los trámites y los datos resultado del análisis de los *logs* del DGM, esta fecha es comparada con la fecha de inserción o modificación de los datos en las bases de datos fuente.

A continuación se muestran ejemplos de ETL realizados:

Capítulo 3. Implementación de la solución

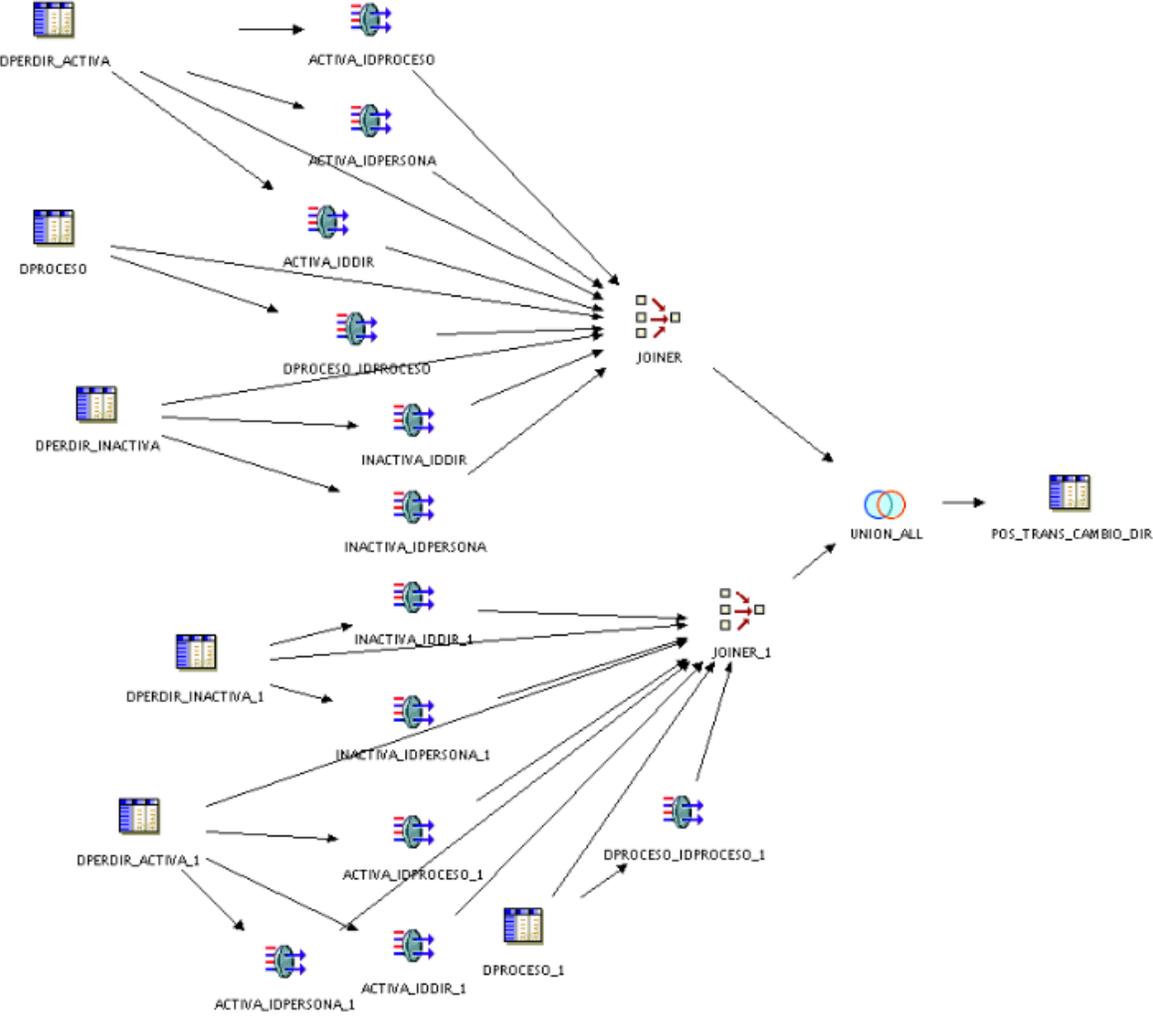


Figura 8. ETL para la carga de los cambios de dirección.

Capítulo 3. Implementación de la solución

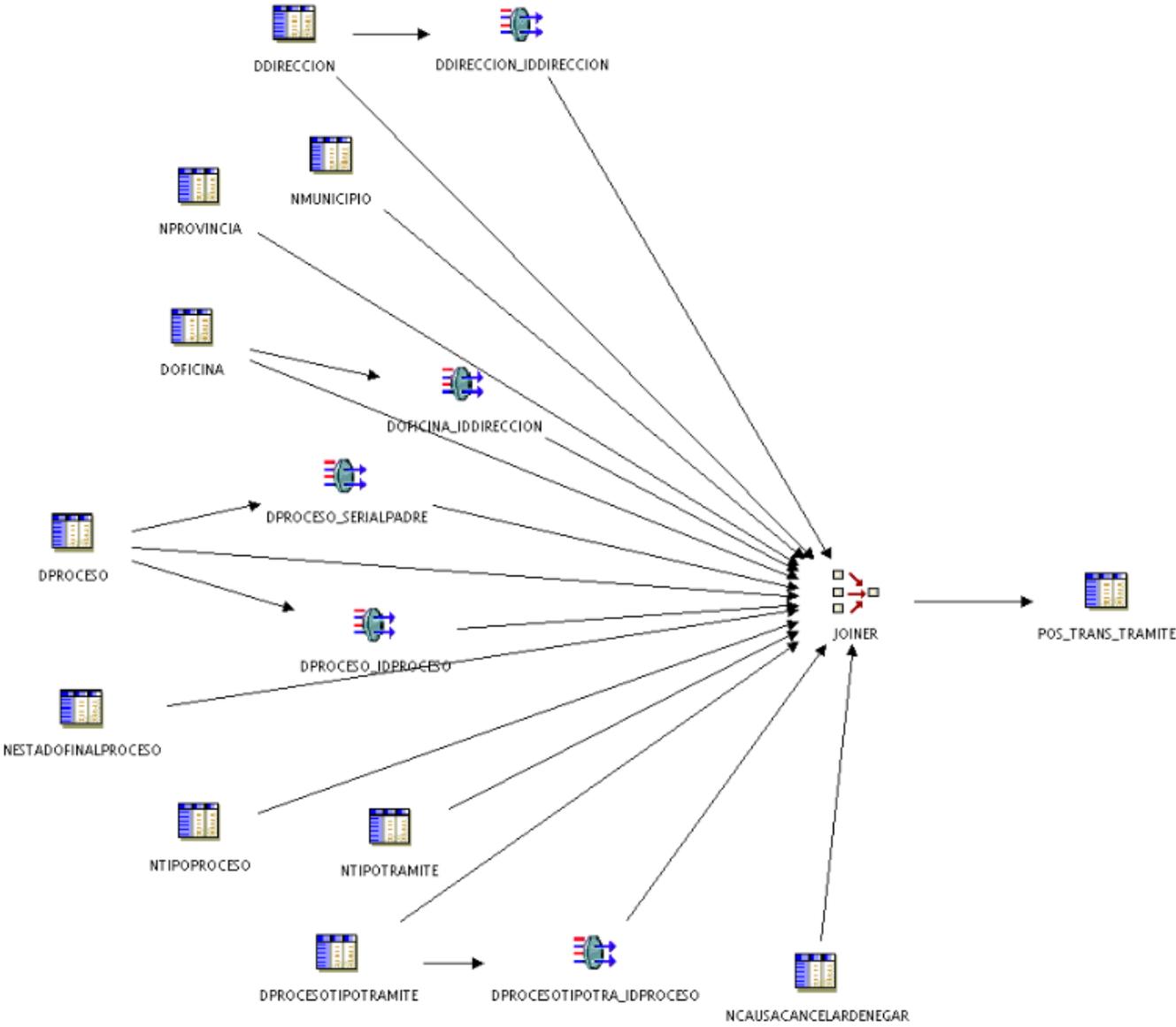


Figura 9. ETL para la carga de los datos relativos a los trámites. Ver el resto de las ETL en el Anexo 4.

Capítulo 3. Implementación de la solución

3.2.2. Implementación de los trabajos (Jobs)

Un trabajo o Job es la realización de un conjunto de tareas para realizar determinadas acciones; también se pueden realizar un grupo de transformaciones dependiendo de la secuencia a seguir, una transformación no se empieza a seguir si la anterior no ha terminado. Estos serán ejecutados diariamente a las 12:00 AM para cargar así los datos correspondientes a ese día de trabajo para no afectar el rendimiento de SUIN.

Cada uno de los trabajos implementados cumple con la siguiente secuencia:

1. Son cargados los datos hacia el área de procesamientos.
2. Son cargados los datos hacia las dimensiones.
3. Son cargados los datos hacia las tablas hecho.

A continuación se presentan ejemplos de trabajos implementados:

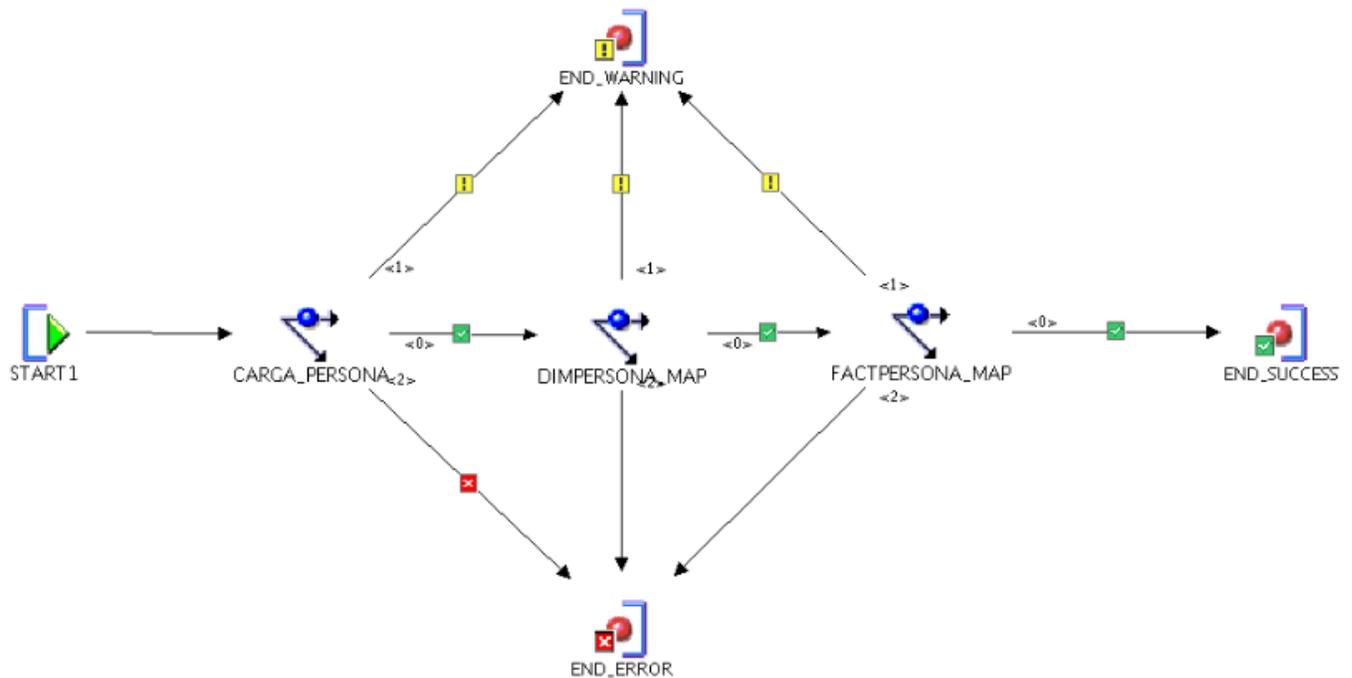


Figura 10. Job para la carga de los datos de las personas.

Capítulo 3. Implementación de la solución

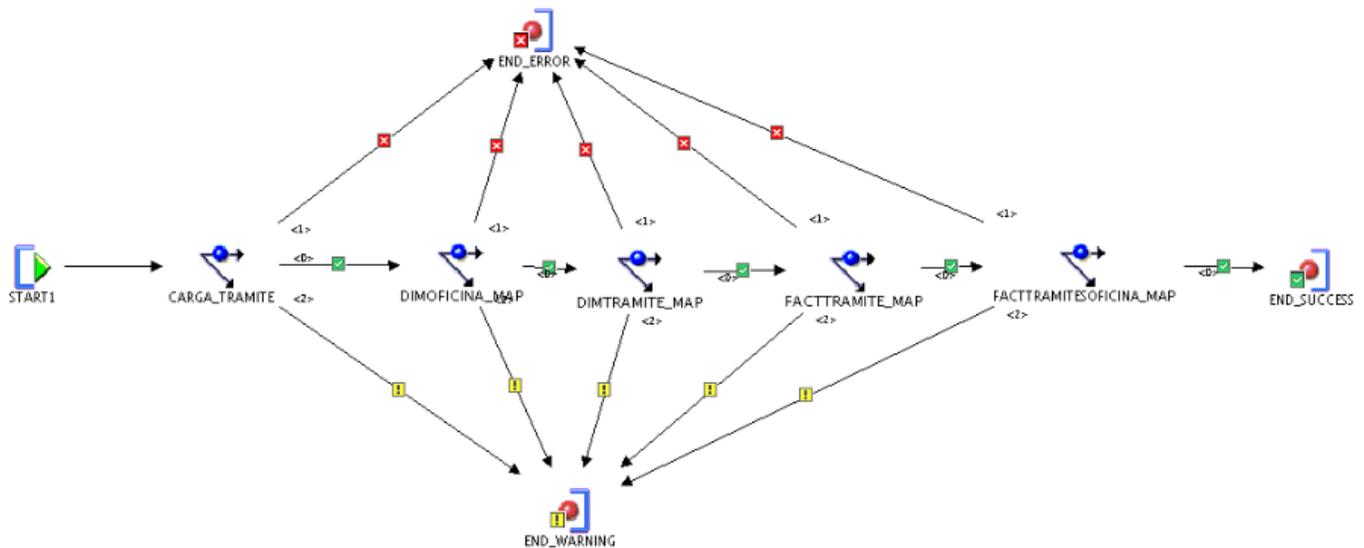


Figura 11. Job para la carga de los datos de los trámites.

3.3. Implementación de los procesos de BI

La inteligencia de negocios es una forma de manejar la información histórica a través de la construcción de almacenes de datos o mercados de datos, con el fin de explotarla y de esta forma realizar análisis que facilitan llevar a cabo una mejor toma de decisiones.

Los procesos ETL son los componentes más importantes y de valor añadido de una infraestructura de BI, pero el éxito de una solución de inteligencia de negocio viene ligado a la capa de presentación la cual depende fundamentalmente de la forma en que se organiza la información estructural, estos son los cubos OLAP, los cuales pueden contener medidas de una o más dimensiones.

3.3.1. Implementación de los cubos OLAP

Uno de los elementos fundamentales donde se organiza la información estructuralmente son los cubos OLAP los cuales pueden poseer más de tres dimensiones por lo que son llamados hipercubos. En el caso concreto de la solución se desarrollaron 5 cubos multidimensionales. Los mismos fueron creados haciendo uso de la herramienta Oracle Business Intelligence Administration Tool que se encuentra dentro de la colección Oracle Business Intelligence Enterprise Edition.

Capítulo 3. Implementación de la solución

Los cubos OLAP se definen uno por cada tabla hecho, debido a que estas tablas son las fuentes de información principal de la organización estructural de los datos, además se observan las tablas dimensiones relacionadas con este hecho así como las medidas del mismo.



Figura 12. Cubos OLAP.

3.3.2. Arquitectura de la información

Los reportes serán mostrados en distintas áreas de análisis para facilitar el acceso a los mismos, agrupando los mismos de acuerdo a la información que muestran.

Estas áreas son:

- Cuadro de Mando Integral: Se muestran los reportes referentes a cada uno de los indicadores brindados que forman parte del cuadro mando.
- Trámites: Se muestran los reportes que muestran información acerca de los trámites realizados.

Ejemplos de estos reportes son:

- Cantidad de trámites por tipo.
- Trámites Realizados por Solicitud de Documentos Actualización.

Capítulo 3. Implementación de la solución

- Población: Se muestran los reportes que muestran información relativa a la población. Ejemplo de estos reportes son:
 - Cantidad de personas por rango de edad.
 - Total de personas por sexo y color de piel.



Figura 13. Mapa de navegación.

3.3.3. Reportes candidatos

A partir de los requisitos de información son definidos una serie de reportes candidatos como finalidad del producto. Usándose para mostrar estos reportes:

- Tablas: Para mostrar los reportes.
- Gráficas: Se usan en conjunto con las tablas.
- Indicador: Mostrar el comportamiento de los indicadores.

A continuación se muestran ejemplos de estos reportes:

Capítulo 3. Implementación de la solución

Cantidad de Trámites por tipo

Provincia	Total	Tipo de Trámite														
		Actualización de Tarjeta Menor	Arribo	Cambio de dirección	Cambio legal	Confección de pasaporte	Defunción	Desmovilización FAR	Deterioro	Entrada país	Nacimiento CIRP	Nacimiento Hospital	Pérdida	Salida prisión	Subsanación de error	
Artemisa	4.034	17	3	300	40	20	112		259	87	9	9	324	29	13	123
Mayabeque	1.408	6	7	79	8	52			100	5	4		29	12	11	35
Total general	5.442	23	10	379	48	72	112		359	92	13	9	353	41	24	158

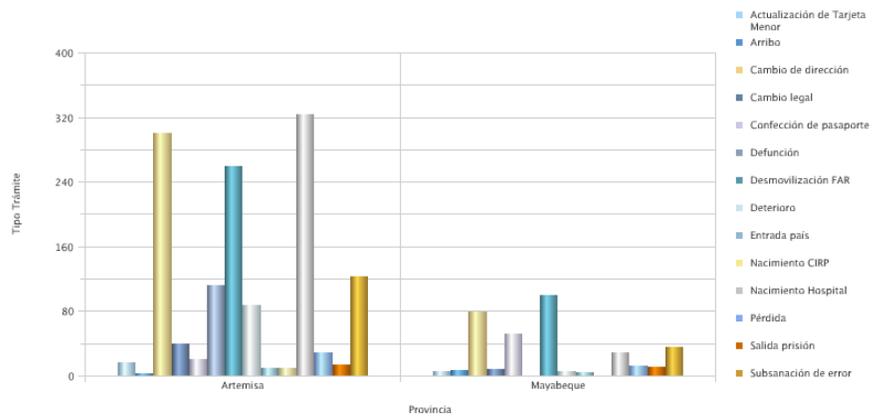


Figura 14. Reporte cantidad por tipo de trámite.

Trámites por Actualización y Solicitud de Documentos

Provincia	Total general	Actualización				Solicitud de Documentos			
		Sub Total	Subsanación de error	Deterioro	Cambio legal	Sub Total	Salida FAR	Salida de Prisiones	Reasentados
Artemisa	4.034	250	123	87	40	281	259	4.034	9
Mayabeque	1.408	48	35	5	8	115	100	1.408	4
Total general	5.442	298	158	92	48	396	359	5.442	13

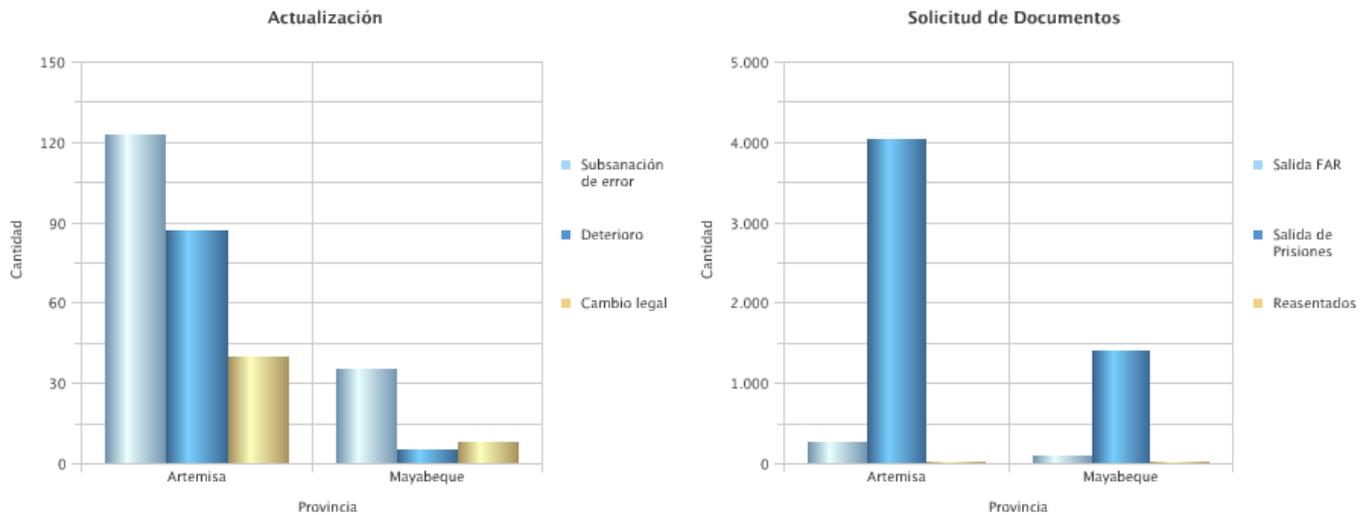


Figura 15. Cantidad de trámites por actualización y solicitud de documentos.

Capítulo 3. Implementación de la solución

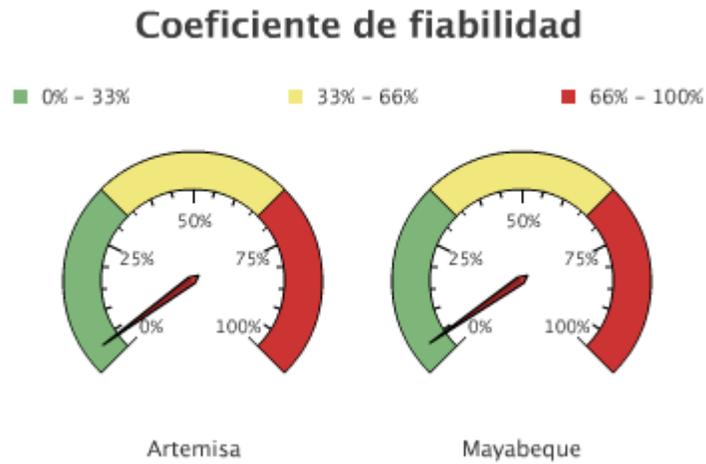


Figura 16. Coeficiente de fiabilidad.

3.4. Diagrama de despliegue

El diagrama de despliegue provee un modelo detallado de la forma en la que los componentes se desplegarán a lo largo de la infraestructura del sistema. Detalla las capacidades de red, las especificaciones del servidor, los requisitos de hardware y otra información relacionada al despliegue del sistema propuesto. (26)

A continuación se muestra el diagrama de despliegue propuesto para la solución.

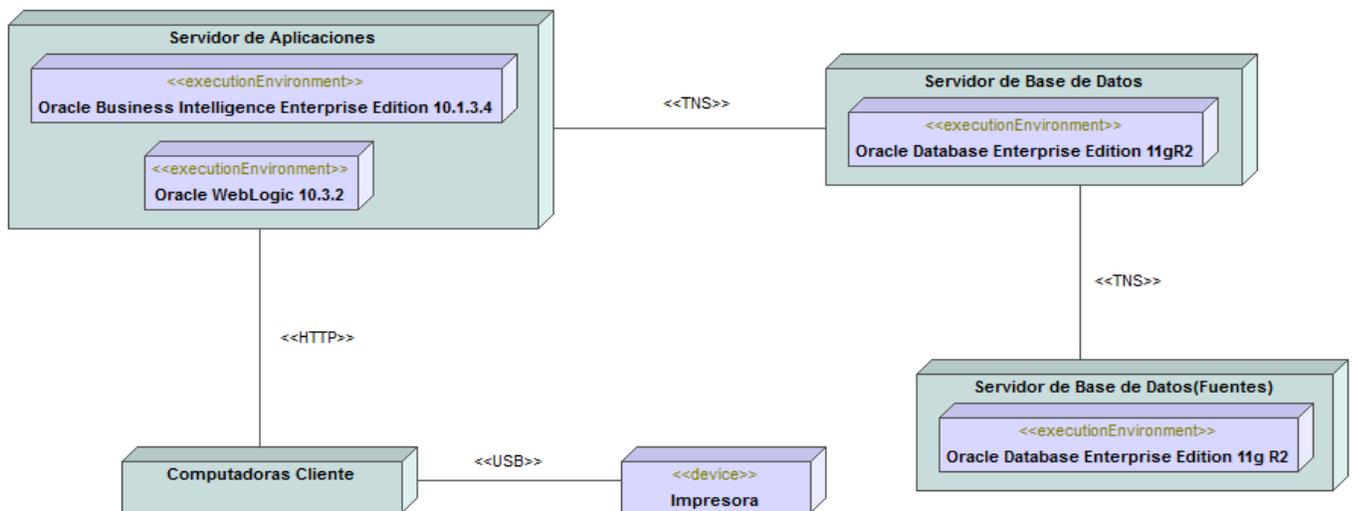


Figura 17. Modelo de despliegue.

Capítulo 3. Implementación de la solución

El diagrama de despliegue propuesto para la solución contiene un servidor de base de datos que posee Oracle Database Enterprise Edition 11gR2 de donde son extraídos los datos a través del protocolo TNS hacia otro servidor de base de datos con las mismas características este se comunica mediante el protocolo TNS con el servidor de aplicaciones el cual debe contar con Oracle Business Intelligence Enterprise Edition 10.1.3.4 y Oracle WebLogic 10.3.2, las computadoras cliente interactúan con el mismo usando el protocolo HTTP, estas a su vez poseen una impresora conectada a través del puerto USB.

3.5. Conclusiones

En este capítulo se realizó la implementación de la base de datos, la cual cuenta con 15 tablas, 10 dimensiones y 5 hechos, 9 secuencias y 38 índices, fueron implementados 19 transformaciones y 3 trabajos para la realización del proceso de ETL; fueron identificados 5 cubos multidimensionales, la arquitectura de la información donde se definieron 3 áreas de análisis, fueron realizados los reportes candidatos, así como el diagrama de despliegue.

Capítulo 4. Validación y pruebas de la Solución de Inteligencia de Negocios

4.1. Introducción

En el presente capítulo se describe el proceso que se realiza para la validación de la Solución de Inteligencia de Negocios para el Sistema de Identificación Nacional de la República de Cuba. Se muestra el resultado de las pruebas de integridad aplicadas sobre los datos, de rendimiento aplicadas sobre el sistema y de aceptación por parte del cliente.

4.2. Validación y pruebas

A nivel mundial se trata de lograr que las empresas que producen software, lo hagan de acuerdo a criterios comunes, para lograr uniformidad. Muchos son los esfuerzos que se realizan para lograr que los productos de software tengan una alta calidad al salir al mercado y por lo tanto que la satisfacción del cliente sea elevada. Diversos autores han mencionados diferentes conceptos relacionados con la calidad de los productos de software, algunos de estos conceptos son los que a continuación se muestran:

“La calidad del software es el grado con el que un sistema, componente o proceso cumple los requerimientos específicos y las necesidades o expectativas del cliente o usuario.”[IEEE, Std 610-1900]

“Concordancia del software producido con los requerimientos explícitamente establecidos, con los estándares de desarrollo prefijados y con los requerimientos implícitos no establecidos formalmente, que desea el usuario.” (27)

La fase de pruebas es la última del ciclo de vida del *software* antes de ser entregado para su explotación y garantizan que el software funcione correctamente y presente la calidad requerida. Las pruebas se realizan para determinar la rapidez con que realiza una tarea bajo condiciones particulares de trabajo. Validan y verifican la calidad del sistema, tales como la escalabilidad y fiabilidad. Verifican el correcto funcionamiento de los pasos de implantación del sistema.

4.3. Tipos de pruebas a realizar

Prueba de integridad de los datos: Permite verificar que los datos mostrados se corresponden con los datos persistentes en la base de datos. Demuestra la veracidad y confiabilidad de la información presentada al usuario.

Prueba de estrés: Se realiza para determinar la solidez de la aplicación en los momentos de carga extrema. Se va aumentando el número de usuarios conectados a la aplicación hasta que colapse,

Capítulo 4. Validación y pruebas de la solución

permitiéndole al administrador del sistema determinar si la aplicación rendirá lo suficiente en caso de que la carga real supere a la carga esperada.

Prueba de carga: Se aplican para determinar el comportamiento de una aplicación bajo una cantidad de peticiones esperada. Esta carga puede ser el número esperado de usuarios concurrentes utilizando la aplicación y que realizan un número específico de transacciones durante el tiempo que dura la carga, de las que se puede mostrar los tiempos de respuestas.

Prueba de aceptación: Son realizadas principalmente por los usuarios con el apoyo del equipo del proyecto. El propósito es confirmar que el sistema está terminado, que desarrolla puntualmente las necesidades de la organización y que es aceptado por los usuarios finales.

4.4. Resultados de la aplicación de las pruebas

4.4.1. Resultados de la prueba de integridad de los datos

Para la realización de las pruebas de integridad fueron definidas 4 iteraciones; de un total de 19 procesos de ETL se tomó una muestra de 6 arrojando como resultado:

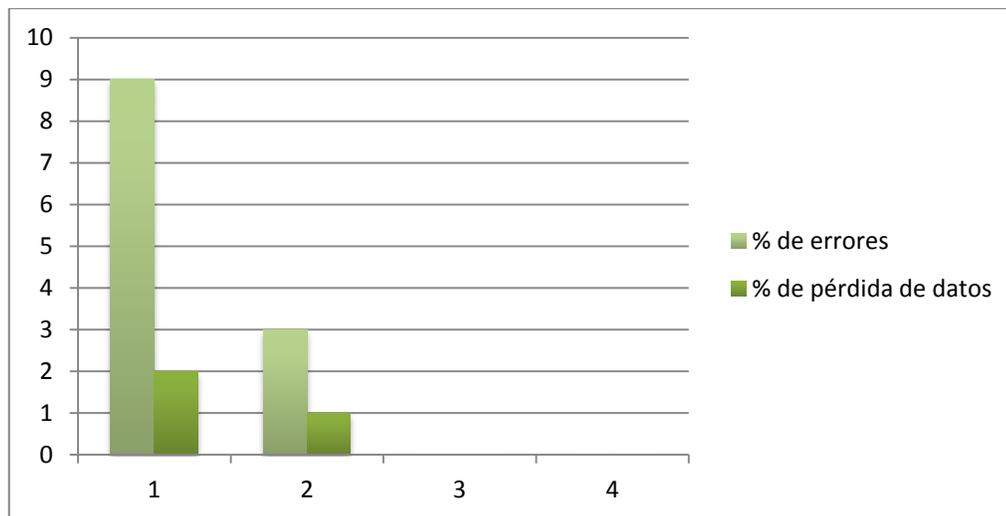


Figura 18. Por ciento de errores por iteración.

A continuación se muestra un ejemplo de la prueba realizada a los datos de los trámites:

Resultado de consulta realizada a la base de datos:

Capítulo 4. Validación y pruebas de la solución

	PROV	TOTAL	ACT_TARJ_MENOF	ARRIBC	CAM_DIF	CAMLEGAL	CONFPAS	DESMOV_FAF	DETERIORC	ENTPAIS	NACHOSF	PERDID	SALPRISIOI	SUBERROR
1	Artemisa ...	4034	17	3	300	40	20	259	87	9	324	29	13	12
2	Mayabequ...	1408	6	7	79	8	52	100	5	4	29	12	11	3

Figura 19. Datos que se encuentran persistidos en la base de datos.

Cantidad de Trámites por tipo

Provincia	Total	Tipo de Trámite													
		Actualización de Tarjeta Menor	Arribo	Cambio de dirección	Cambio legal	Confección de pasaporte	Defunción	Desmovilización FAR	Deterioro	Entrada país	Nacimiento CIRP	Nacimiento Hospital	Pérdida	Salida prisión	Subsanación de error
Artemisa	4.034	17	3	300	40	20	112	259	87	9	9	324	29	13	123
Mayabeque	1.408	6	7	79	8	52		100	5	4		29	12	11	35
Total general	5.442	23	10	379	48	72	112	359	92	13	9	353	41	24	158

Figura 20. Reporte mostrado en la aplicación.

Se puede concluir que no se produce pérdida de los datos al realizarse los procesos de extracción, transformación y carga desde las fuentes de datos.

4.4.2. Resultados de la prueba de carga y estrés

Estas pruebas fueron realizadas sobre un ordenador con 2 Gb de RAM y un procesador Intel Core 2 Duo a 2.0 GHz; para la realización de las pruebas de carga y estrés se fueron aumentando la cantidad de usuarios de 50 en 50 hasta llegar al momento donde el sistema comienza responder con errores lo cual sucede al cabo de 300 usuarios, mostrándose errores en un 4% de las peticiones.

Capítulo 4. Validación y pruebas de la solución

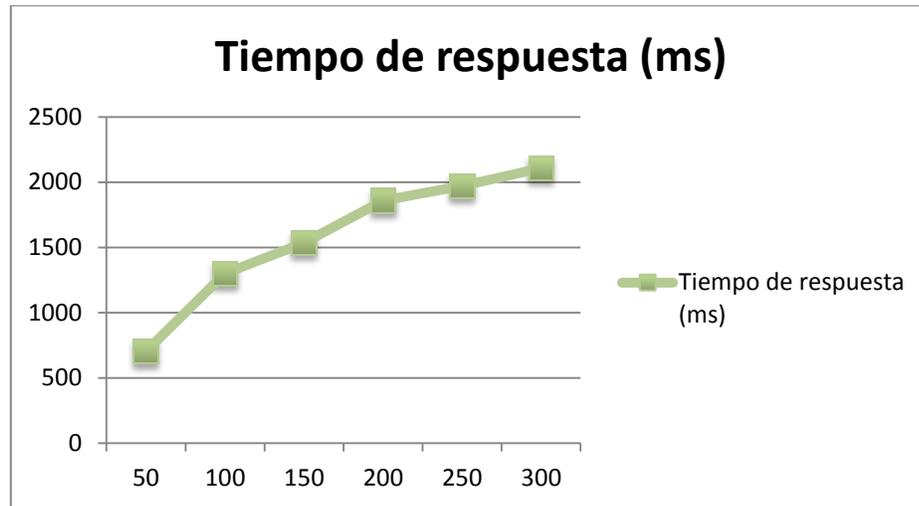


Figura 21. Muestra el tiempo medio de respuesta por la cantidad de usuarios.

Por lo que podemos concluir que sobre un hardware de 2Gb de RAM y un procesador Intel Core 2 Duo, la aplicación responde correctamente ante 250 peticiones concurrentes, lo cual supera la cantidad posible de usuarios que posee la aplicación.

4.4.3. Resultados de la prueba de aceptación

Fueron realizadas las pruebas de aceptación por parte de CALISOFT en conjunto con el cliente, aceptándose la solución.

Acta de Aceptación



Proyecto: Modernización del sistema de Identificación, Inmigración y Extranjería de Cuba (Identidad Cuba).

Producto: Sistema Único de Identificación Nacional de la Población de la República de Cuba (SUIN).

Categoría de las pruebas: Revisión a la aplicación.

Fecha de conciliación: 22 de Mayo de 2012

Observaciones del proceso:

Durante el proceso de pruebas de aceptación se detectaron un conjunto de No Conformidades (NC) y Pedidos de Cambio (PC) que fueron registrados adecuadamente en los correspondientes informes diarios de NC y PC, con sus respectivas observaciones.

Teniendo en cuenta que fue validada por el cliente la corrección de la mayoría de las No Conformidades y gestionados los Pedidos de Cambios y valorando que las No Conformidades pendientes de solución no se consideran críticas por parte del cliente, para la etapa de despliegue del sistema, se ha tomado el acuerdo de Aceptar la aplicación Sistema Único de Identificación Nacional de la Población de la República de Cuba (SUIN) en su versión 1.0 y el documento "Manual de Usuario" en su versión 1.0, todos con fecha 21 de Mayo de 2012.

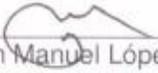
El conjunto de incidencias pendientes, de las que fueron detectadas durante el proceso de pruebas de aceptación han sido especificadas en los documentos Informe Final de No Conformidades (NC) e Informe Final de Pedidos de Cambio (PC), con sus respectivas observaciones.

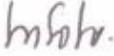
Capítulo 4. Validación y pruebas de la solución

Acta de Aceptación



Para que conste la Aceptación de los resultados de las pruebas y por tanto la Aceptación de los entregables especificados, dando fe del acuerdo, se extiende la presente **Acta** en tres (3) ejemplares, rubricados por los principales **Representantes de las Partes**.


Coronel Juan Manuel López Acevedo
Jefe de Dirección de Identificación y Registros


Tte Coronel Magnolia Soto Bernal
Jefe de Departamento CIRP


MsC. Yancy Capote García
Representante Centro Nacional de Calidad de Software (CALISOFT)

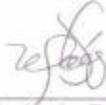

MsC. Yudenia Ramírez Mastrapa
Jefa del proyecto



Figura 22. Carta de aceptación.

4.5. Análisis de la rapidez de obtención de los reportes

De un total de 14 reportes fue tomada una muestra de 7 para el análisis del tiempo que demora la obtención de los datos en los reportes que se muestran el SUIN y en la solución lo cual arrojó los siguientes resultados:

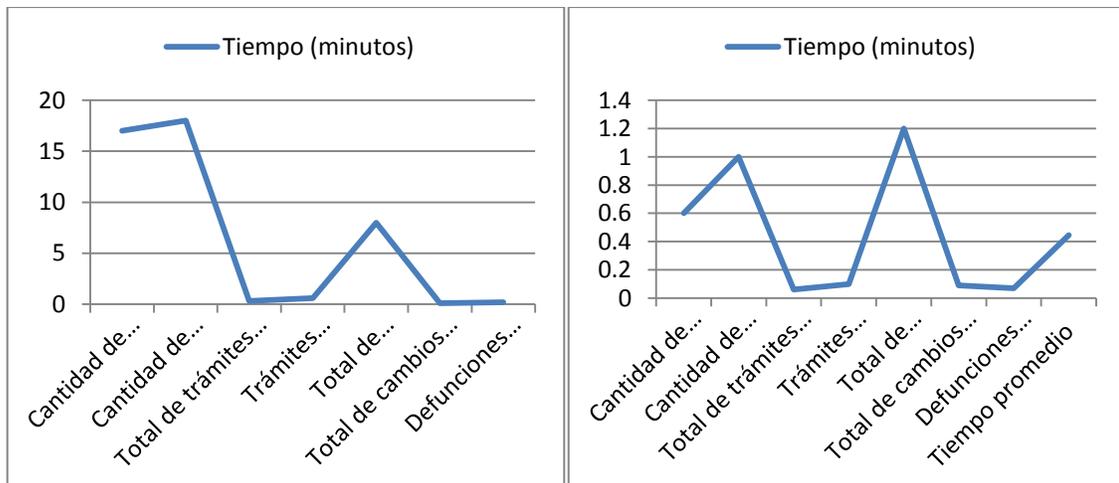
Capítulo 4. Validación y pruebas de la solución

Reportes	Tiempo (minutos)
Cantidad de personas por sexo	17
Cantidad de personas por rango de edad	18
Total de trámites realizados	0.3
Trámites realizados por tipo	0.6
Total de residentes por color de piel	8
Total de cambios de direcciones	0.1
Defunciones realizadas	0.2
Tiempo promedio	6.314285714

Tabla 8. Tiempo de obtención de los reportes del SUIN.

Reportes	Tiempo (minutos)
Cantidad de personas por sexo	0.6
Cantidad de personas por rango de edad	1
Total de trámites realizados	0.06
Trámites realizados por tipo	0.1
Total de residentes por color de piel	1.2
Total de cambios de direcciones	0.09
Defunciones realizadas	0.07
Tiempo promedio	0.445714286

Tabla 9. Tiempo de obtención de los reportes por la solución implementada.



Capítulo 4. Validación y pruebas de la solución

Comparación de los tiempos de obtención de reportes por el SUIN (gráfica izquierda) y la solución implementada demostrándose el mejoramiento en los tiempos de respuesta por parte de la solución implementada.

4.6. Beneficios de la aplicación

La implementación de una solución de inteligencia de negocio permite a los funcionarios del MININT:

- Recolectar y procesar la información generada a partir de los trámites realizados en la DIR.
- Facilita la gestión histórica.
- Proporciona acceso a información clave para la toma de decisiones.
- Permite la realización de reportes personalizados de manera rápida y sencilla simplificando así la colaboración y su uso compartido, dando una visión profunda del comportamiento de los indicadores organizacionales.
- Ofrece una capa de servicios SOAP que permita a aplicaciones externas tener acceso a los reportes creados.

4.7. Conclusiones

Con la realización de las pruebas definidas se comprobó el desempeño de la aplicación y los resultados fueron los esperados. La prueba de integridad de los datos mostrados en los reportes con respecto a los datos persistentes en la base de datos arrojó excelentes resultados, confirmándose que no se produce pérdida de los datos al realizarse los procesos de extracción, transformación y carga desde las fuentes de datos. A través de las pruebas de carga y estrés, se pudo probar que la aplicación tiene un buen funcionamiento con el número de usuarios concurrentes que se espera accedan a la misma. Fueron realizadas pruebas comparando los tiempos en que se muestran los reportes en el SUIN y en la solución implementada quedando demostrada la reducción de los tiempos de obtención de los reportes.

Conclusiones

Conclusiones

La investigación realizada sobre la gestión de la información histórica y generación de reportes y el desarrollo de la Solución de Inteligencia de Negocios para el SUIN disminuyéndose los tiempos de obtención de los reportes de un promedio de 6 minutos a solo 26 segundos permitiendo a los autores concluir que:

- El estudio de los principales conceptos asociados a las soluciones de inteligencia de negocios, metodologías y herramientas a utilizar, mostró la necesidad de agilizar el proceso de desarrollo a través del uso de la metodología Kimball, las herramientas permitieron realizar una correcta elaboración de la solución, a través de la explotación de las ventajas que estas ofrecen.
- El estudio, análisis del negocio y las necesidades del usuario posibilitó la identificación de 8 requisitos de información, 7 requisitos funcionales y 24 no funcionales.
- Fueron definidos los hechos y dimensiones que conforman la matriz bus, así como el diseño del modelo de datos para el *data warehouse*, de la arquitectura de la solución y la arquitectura de la información.
- 19 procesos de ETL permitieron realizar la extracción de los datos desde las fuentes, procesarlos y cargarlos hacia el almacén de datos.
- La creación de la aplicación garantiza la gestión de la información histórica y la generación de reportes facilitando el acceso a la información para el apoyo a la toma de decisiones por parte de los funcionarios del MININT.

Referencias bibliográficas

Trabajos citados

1. **Emprendimientos Corporativos S.A.** iProfesional.com. *iProfesional.com*. [En línea] 2010. [Citado el: 12 de diciembre de 2011.] <http://www.iprofesional.com/notas/51865-Los-beneficios-de-la-inteligencia-de-negocios.html>.
2. **Identidad Cuba.UCI.** *Proyecto Técnico.DIR*. 2009.
3. **Proyecto Identidad Cuba.UCI.** *Arquitectura de Software*. [Documento] 2010.
4. **The Data Warehousing Institute.** tdwi. [En línea] [Citado el: 10 de Diciembre de 2011.]
5. **DB System.** DB System. *DB System*. [En línea] 2008. [Citado el: 10 de 12 de 2011.] http://www.db-system.com/portal/page?_pageid=33,7665&_dad=portal&_schema=PORTAL.
6. **Inteligencia de Negocios S.A.** in. *in*. [En línea] 2011. [Citado el: 10 de 12 de 2011.] <http://www.idensa.com/>.
7. **Universidad de Holguín.** Plataforma de Aprendizaje de la Universidad de Holguín. *Plataforma de Aprendizaje de la Universidad de Holguín*. [En línea] 2011. [Citado el: 12 de 11 de 2011.] http://www.google.com.cu/url?sa=t&rct=j&q=base+de+datos&source=web&cd=17&ved=0CHoQFjAGOAo&url=http%3A%2F%2Fmoodle.uho.edu.cu%2Fmod%2Fresource%2Fview.php%3Finpopup%3Dtrue%26id%3D3528&ei=e_K6T5TRGoPqgAfa0qXnCc&usq=AFQjCNGRBL--sywqToaJjf426okhnuX3-w&cad=rja.
8. **Escobar Domínguez, René Raydel.** NETRONICS. *NETRONICS*. [En línea] 2011. [Citado el: 12 de 11 de 2011.] http://www.netronics.com/clasificacion_de_base_datos.html.
9. **Sinnexus.** Sinnexus. *Sinnexus*. [En línea] 2011. [Citado el: 13 de 12 de 2011.] http://www.sinnexus.com/business_intelligence/olap_avanzado.aspx.
10. **Imhof, Claudia, Galemmo, Nicholas y G. Geiger, Jonathan.** *Mastering Data Warehouse Design Relational and Dimensional Techniques*.
11. **Padrón Torres, Liudmila.** mailxmailx. *mailxmailx*. [En línea] 17 de 11 de 2010. [Citado el: 11 de 12 de 2011.] <http://www.mailxmail.com/curso-almacenes-datos-importancia-estandar/caracteristicas-almacenes-datos>.
12. **Clemente García, Gerardo.** Riunet. *Riunet*. [En línea] 2008. [Citado el: 11 de 12 de 2011.] <http://riunet.upv.es/manakin/bitstream/handle/10251/2505/tesisUPV2842.pdf>.
13. **Berzal, Fernando.** [En línea] <http://elvex.ugr.es/idbis/db/docs/intro/F%20Modelo%20multidimensional.pdf>.

Referencias bibliográficas

14. **Gloria Wolff, Carmen.** Modelamiento multidimensional. *Modelamiento multidimensional*. [En línea] 28 de 8 de 2002. [Citado el: 12 de 12 de 2011.] <http://www.inf.udec.cl/~revista/ediciones/edicion4/modmulti.PDF>.
15. **Valencia Arcos, Janeth del Carmen y Guevara Lenis, Jorge Eduardo.** [En línea] 6 de 2007. [Citado el: 12 de 12 de 2011.] <http://bibdigital.epn.edu.ec/bitstream/15000/445/1/CD-0827.pdf>.
16. **Casales Cabrera, María Evilia.** Facultad de Ciencias. Universidad Nacional Autónoma de México. *Facultad de Ciencias. Universidad Nacional Autónoma de México*. [En línea] 2009. [Citado el: 21 de noviembre de 2011.] <http://hp.fciencias.unam.mx/~alg/bd/dwh.pdf>.
17. **Límia Navarro, Alberto, y otros, y otros.** *Metodología de Desarrollo de Soluciones BI y Warehousing(BI & W)*. Ciudad de La Habana, Universidad de las Ciencias Informáticas : s.n.
18. **Embarcadero.** E/R Studio Modelado de datos empresarial.Database Tools and Developer Software. [En línea] [Citado el: 10 de enero de 2011.] <http://www.embarcadero.com>.
19. **Oracle Corporation.** Oracle. [En línea] 2010. [Citado el: 12 de enero de 2011.] <http://www.oracle.com/technology/products/database/oracle11g>.
20. **Bob Bryla, Kevin Loney.** *Oracle Database 11g DBA Handbook (Osborne ORACLE Press Series)*.
21. **Oracle Corporation.** PL / SQL. [En línea] 2010. [Citado el: 11 de enero de 2011.] http://www.oracle.com/technology/tech/pl_sql.
22. **Torres Miyares, Daniel Ramón y Silva Fernández, Yanet.** *Tesis Common Services*. 2011.
23. **Identidad Cuba.** Modelo de datos Identidad. [En línea] 2011. [Citado el: 12 de 12 de 2011.] <http://intranet.uci.dev/bdsite/Shared%20Documents/Release/Identidad/ID-BD1011-2012-Modelo%20de%20Datos-IDENTIDAD.dm1>.
24. —. Modelo de datos ADM. [En línea] 2011. [Citado el: 12 de 12 de 2011.] <http://intranet.uci.dev/bdsite/Shared%20Documents/Release/Administraci%C3%B3n/ADM-BD101-2012-ADM-Modelo%20de%20Datos-ADM.dm1>.
25. **Microsoft.** MSDN. *MSDN*. [En línea] 2012. [Citado el: 5 de 4 de 2012.] <http://msdn.microsoft.com/es-es/library/ms190457.aspx>.
26. **Sparxsystems.** Enterprise Architect. [En línea] 2011. [Citado el: 14 de 1 de 2012.] http://www.sparxsystems.com.ar/resources/tutorial/physical_models.html.
27. **Pressman, Roger.** *Ingeniería del Software: Un Enfoque Práctico*. s.l. : McGraw-Hill, 2006.

Referencias bibliográficas

28. **Gravitar.** Gravitar Información sin Límites. *Gravitar*. [En línea] 2011. [Citado el: 12 de diciembre de 2011.] <http://www.gravitar.biz/index.php/herramientas-bi/pentaho/>.
29. **W3 Schools.** W3 Schools. [En línea] 2010. <http://www.w3schools.com/soap/default.asp>.

Anexos

Anexo 1. Descripción de los requisitos funcionales

RF 1. Extraer, transformar y cargar datos de los sistemas fuentes.

Propósito	Permite cargar datos de los sistemas fuentes hacia el almacén de datos.	
Roles	Sistema	
Precondiciones	Debe existir conectividad directa entre los servidores de BD origen y destino.	
Entidades tratadas	Entidad <ul style="list-style-type: none">• DDISPOSITIVOSLOGS• DESTTRABAGRUPLOGICA• POS_TRANS_DGM• DDISPOSITIVO• POS_TRANS_TRAMITE• DPERDIR_ACTIVIA• DPERDIR_INACTIVA• DPROCESO• POS_TRANS_CAMBIO_DIR• DDIRECCION• NPROVINCIA• NMUNICIPIO• NNIVELCULTURAL• NCOLORCABELLO• NCOLOROJOS• NCOLORPIEL• DPERSONADIRECCION• DPERSONAOTROSDATOS• NCONDICIONMIGRATORIA• DNACIONAL	Atributos Ver modelo de datos de los esquemas Identidad, ADM y DGM.

Anexos

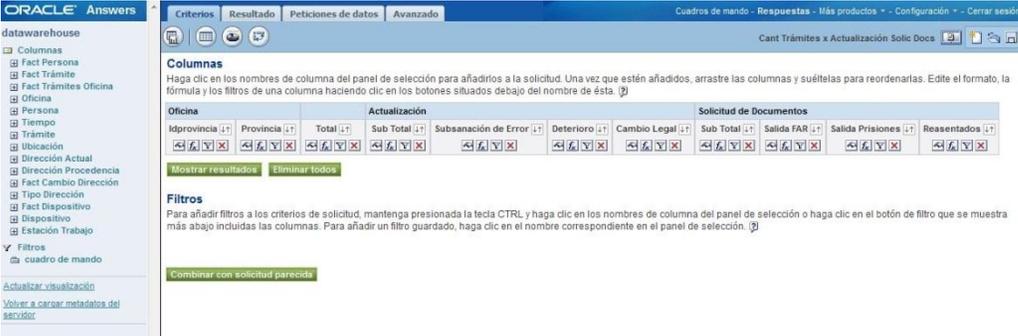
	<ul style="list-style-type: none"> • DPERSONA • NTIPOPERSONA • POS_TRANS_PERSONA • NTIPOPROCESO • DOFICINA • NCAUSACANCELARDENEGAR • NTIPOTRAMITE • NESTADOFINALPROCESO • DPROCESOTIPOTRAMITE 	
Descripción	<p>1.1 Se cargan los datos en horario nocturno desde los esquemas ADM, Identidad y de los logs generados por el DGM hacia el área de procesamiento del almacén de datos.</p> <p>1.2 Se cargan los datos hacia las dimensiones.</p> <p>1.3 Se cargan los datos hacia las tablas de hechos.</p>	
Validaciones	<p>1. Las medidas de las tablas hechos no pueden tener valores nulos o negativos.</p>	
Poscondiciones	<p>1. El almacén de datos debe incluir los datos almacenados en los sistemas fuentes desde la última carga realizada.</p>	
Prototipo		

RF 2. Mostrar reporte.

Propósito	Permite mostrar los reportes solicitados por los usuarios.
Roles	Administrador, Administrador de Reportes, Analista de Información
Precondiciones	<ol style="list-style-type: none"> 1. Los usuarios deben estar autenticados en el sistema. 2. Debe existir al menos un reporte al que el usuario tenga permisos de

	lectura.																																																																																																																																																	
Entidades tratadas	Entidad	Atributos																																																																																																																																																
Descripción	2.1 Seleccionar el área de información. 2.2 Seleccionar la página correspondiente al reporte que se desee visualizar.																																																																																																																																																	
Validaciones																																																																																																																																																		
Poscondiciones	Se muestra el reporte seleccionado.																																																																																																																																																	
Prototipo	<p>Cantidad de personas por rangos de edades</p> <table border="1"> <thead> <tr> <th>Provincia</th> <th>Total Residentes</th> <th>MENORES DE 16 AÑOS</th> <th>DE 16 A 25 AÑOS</th> <th>DE 26 A 35 AÑOS</th> <th>DE 36 A 45 AÑOS</th> <th>DE 46 A 55 AÑOS</th> <th>DE 56 A 65 AÑOS</th> </tr> </thead> <tbody> <tr><td>Pinar Del Rio</td><td>598.847</td><td>101.111</td><td>82.088</td><td>72.320</td><td>101.471</td><td>91.201</td><td>67.197</td></tr> <tr><td>Artemisa</td><td>522.215</td><td>83.846</td><td>71.406</td><td>64.876</td><td>90.806</td><td>81.073</td><td>53.638</td></tr> <tr><td>La Habana</td><td>2.219.533</td><td>320.403</td><td>279.539</td><td>248.316</td><td>363.424</td><td>376.775</td><td>254.199</td></tr> <tr><td>Mayabeque</td><td>397.462</td><td>59.684</td><td>51.704</td><td>47.534</td><td>66.611</td><td>60.974</td><td>43.133</td></tr> <tr><td>Matanzas</td><td>702.889</td><td>108.211</td><td>92.458</td><td>85.689</td><td>126.949</td><td>111.401</td><td>77.137</td></tr> <tr><td>Villa Clara</td><td>792.199</td><td>121.314</td><td>102.043</td><td>89.227</td><td>134.923</td><td>122.634</td><td>91.076</td></tr> <tr><td>Cienfuegos</td><td>399.573</td><td>65.691</td><td>53.581</td><td>48.654</td><td>70.181</td><td>63.802</td><td>42.714</td></tr> <tr><td>Sancti Spiritus</td><td>462.849</td><td>76.262</td><td>60.087</td><td>55.230</td><td>84.734</td><td>70.275</td><td>50.487</td></tr> <tr><td>Ciego De Avila</td><td>454.990</td><td>72.744</td><td>62.037</td><td>56.902</td><td>80.632</td><td>66.763</td><td>45.900</td></tr> <tr><td>Camaguey</td><td>834.224</td><td>122.922</td><td>104.237</td><td>101.865</td><td>146.453</td><td>129.139</td><td>90.508</td></tr> <tr><td>Las Tunas</td><td>553.398</td><td>92.184</td><td>76.206</td><td>71.003</td><td>103.153</td><td>83.726</td><td>58.474</td></tr> <tr><td>Holguin</td><td>1.099.276</td><td>187.158</td><td>140.733</td><td>131.274</td><td>206.871</td><td>165.814</td><td>113.028</td></tr> <tr><td>Granma</td><td>885.817</td><td>154.063</td><td>120.940</td><td>110.444</td><td>163.420</td><td>133.218</td><td>89.292</td></tr> <tr><td>Santiago Cuba</td><td>1.076.400</td><td>189.902</td><td>152.308</td><td>141.590</td><td>194.295</td><td>159.076</td><td>110.726</td></tr> <tr><td>Guantanamo</td><td>556.268</td><td>102.588</td><td>79.122</td><td>71.288</td><td>99.422</td><td>81.954</td><td>53.924</td></tr> <tr><td>Isla Juventud</td><td>91.107</td><td>15.669</td><td>12.723</td><td>12.008</td><td>16.874</td><td>14.378</td><td>10.192</td></tr> <tr><td>Total general</td><td>11.647.047</td><td>1.873.752</td><td>1.541.212</td><td>1.408.220</td><td>2.050.219</td><td>1.812.203</td><td>1.251.625</td></tr> </tbody> </table> <p style="text-align: center;">Por rangos de edades</p>		Provincia	Total Residentes	MENORES DE 16 AÑOS	DE 16 A 25 AÑOS	DE 26 A 35 AÑOS	DE 36 A 45 AÑOS	DE 46 A 55 AÑOS	DE 56 A 65 AÑOS	Pinar Del Rio	598.847	101.111	82.088	72.320	101.471	91.201	67.197	Artemisa	522.215	83.846	71.406	64.876	90.806	81.073	53.638	La Habana	2.219.533	320.403	279.539	248.316	363.424	376.775	254.199	Mayabeque	397.462	59.684	51.704	47.534	66.611	60.974	43.133	Matanzas	702.889	108.211	92.458	85.689	126.949	111.401	77.137	Villa Clara	792.199	121.314	102.043	89.227	134.923	122.634	91.076	Cienfuegos	399.573	65.691	53.581	48.654	70.181	63.802	42.714	Sancti Spiritus	462.849	76.262	60.087	55.230	84.734	70.275	50.487	Ciego De Avila	454.990	72.744	62.037	56.902	80.632	66.763	45.900	Camaguey	834.224	122.922	104.237	101.865	146.453	129.139	90.508	Las Tunas	553.398	92.184	76.206	71.003	103.153	83.726	58.474	Holguin	1.099.276	187.158	140.733	131.274	206.871	165.814	113.028	Granma	885.817	154.063	120.940	110.444	163.420	133.218	89.292	Santiago Cuba	1.076.400	189.902	152.308	141.590	194.295	159.076	110.726	Guantanamo	556.268	102.588	79.122	71.288	99.422	81.954	53.924	Isla Juventud	91.107	15.669	12.723	12.008	16.874	14.378	10.192	Total general	11.647.047	1.873.752	1.541.212	1.408.220	2.050.219	1.812.203	1.251.625
Provincia	Total Residentes	MENORES DE 16 AÑOS	DE 16 A 25 AÑOS	DE 26 A 35 AÑOS	DE 36 A 45 AÑOS	DE 46 A 55 AÑOS	DE 56 A 65 AÑOS																																																																																																																																											
Pinar Del Rio	598.847	101.111	82.088	72.320	101.471	91.201	67.197																																																																																																																																											
Artemisa	522.215	83.846	71.406	64.876	90.806	81.073	53.638																																																																																																																																											
La Habana	2.219.533	320.403	279.539	248.316	363.424	376.775	254.199																																																																																																																																											
Mayabeque	397.462	59.684	51.704	47.534	66.611	60.974	43.133																																																																																																																																											
Matanzas	702.889	108.211	92.458	85.689	126.949	111.401	77.137																																																																																																																																											
Villa Clara	792.199	121.314	102.043	89.227	134.923	122.634	91.076																																																																																																																																											
Cienfuegos	399.573	65.691	53.581	48.654	70.181	63.802	42.714																																																																																																																																											
Sancti Spiritus	462.849	76.262	60.087	55.230	84.734	70.275	50.487																																																																																																																																											
Ciego De Avila	454.990	72.744	62.037	56.902	80.632	66.763	45.900																																																																																																																																											
Camaguey	834.224	122.922	104.237	101.865	146.453	129.139	90.508																																																																																																																																											
Las Tunas	553.398	92.184	76.206	71.003	103.153	83.726	58.474																																																																																																																																											
Holguin	1.099.276	187.158	140.733	131.274	206.871	165.814	113.028																																																																																																																																											
Granma	885.817	154.063	120.940	110.444	163.420	133.218	89.292																																																																																																																																											
Santiago Cuba	1.076.400	189.902	152.308	141.590	194.295	159.076	110.726																																																																																																																																											
Guantanamo	556.268	102.588	79.122	71.288	99.422	81.954	53.924																																																																																																																																											
Isla Juventud	91.107	15.669	12.723	12.008	16.874	14.378	10.192																																																																																																																																											
Total general	11.647.047	1.873.752	1.541.212	1.408.220	2.050.219	1.812.203	1.251.625																																																																																																																																											

RF 3. Adicionar reporte.

Propósito	Permite agregar un nuevo reporte al sistema de BI.	
Roles	Administrador, Administrador de Reportes	
Precondiciones	<ol style="list-style-type: none"> 1. El usuario debe estar autenticado en el sistema. 2. Debe existir al menos un área temática disponible. 	
Entidades tratadas	Entidad	Atributos
Descripción	<ol style="list-style-type: none"> 3.1 Seleccionar el vínculo "Respuestas". 3.2 Seleccionar el área temática a usar. 3.3 Confeccionar el reporte a partir de los criterios presentes en el área temática seleccionada. 3.4 Guardar el reporte con el nombre que se desee. 	
Validaciones		
Postcondiciones	Debe quedar almacenado en los catálogos de presentación el nuevo reporte.	
Prototipo		

RF 4. Eliminar reporte.

Propósito	Permite eliminar un reporte sistema.
------------------	--------------------------------------

Roles	Administrador, Administrador de Reportes	
Precondiciones	<ol style="list-style-type: none"> 1. El usuario debe estar autenticado en el sistema. 2. Debe existir al menos un reporte almacenado en el sistema. 	
Entidades tratadas	Entidad	Atributos
Descripción	<ol style="list-style-type: none"> 4.1 Seleccionar el vínculo “Configuración”. 4.2 Seleccionar la opción “Administración”. 4.3 Seleccionar “Administrar Catálogo de presentación” de la lista de actividades que se pueden realizar. 4.4 Buscar el reporte a eliminar en la carpeta que se guardó. 4.5 Seleccionar la opción “Eliminar” correspondiente al reporte que se desea eliminar del sistema. 4.6 Se muestra un mensaje de confirmación. 4.7 Seleccionar “Sí” o “No” en caso que se desee. 	
Validaciones		
Postcondiciones	Debe quedar eliminado de los catálogos de presentación el reporte en caso de seleccionar “Sí”.	
Prototipo	 <p>The screenshot shows the 'Administrar catálogo' interface. At the top, there is a search bar and a 'Terminado' button. Below that, there is a section for 'Carpetas compartidas' with a 'cuadro de mando' (command panel). The main area displays a list of folders and reports. The 'Eliminar' button is highlighted. Below the list, there is a section for archiving the catalog content with options to maintain permissions and dates.</p>	

RF 5. Modificar reporte.

Propósito	Permite modificar un reporte existente.	
Roles	Administrador, Administrador de Reportes	
Precondiciones	<ol style="list-style-type: none"> 1. El usuario debe estar autenticado en el sistema. 2. Debe existir al menos un reporte almacenado en el sistema. 	
Entidades tratadas	Entidad	Atributos
Descripción	<ol style="list-style-type: none"> 5.1 Seleccionar el vínculo "Respuestas". 5.2 Seleccionar del catálogo el reporte a modificar. 5.3 Seleccionar la opción "Modificar". 5.4 Modificar el reporte seleccionado. 5.5 Guardas los cambios realizados. 	
Validaciones		
Postcondiciones	Debe quedar modificado el reporte.	
Prototipo		

RF 6. Exportar reporte a PDF.

Propósito	Permite exportar los reportes mostrados en formato PDF.	
Roles	Administrador, Administrador de Reportes, Analista de información	
Precondiciones	<ol style="list-style-type: none"> 1. El usuario debe estar autenticado en el sistema. 2. Debe existir al menos un reporte en el sistema al que el usuario tenga permisos de visualizar. 	
Entidades tratadas	Entidad	Atributos
Descripción	6.1 Seleccionar el ícono “Apto para impresión”. 6.2 Seleccionar la opción “PDF”.	
Validaciones		
Postcondiciones	El sistema debe generar un archivo con formato PDF.	
Prototipo		

RF 7. Exportar reporte a Excel.

Propósito	Permite exportar los reportes mostrados en formato XLS.																																																																																																																															
Roles	Administrador, Administrador de Reportes																																																																																																																															
Precondiciones	<ol style="list-style-type: none"> 1. El usuario debe estar autenticado en el sistema. 2. Debe existir al menos un reporte almacenado en el sistema. 																																																																																																																															
Entidades tratadas	Entidad	Atributos																																																																																																																														
Descripción	<ol style="list-style-type: none"> 7.1 Seleccionar el vínculo "Respuestas". 7.2 Seleccionar del catálogo el reporte deseado. 7.3 Seleccionar el vínculo "Descargar". 7.4 Seleccionar la opción "Descargar en Excel" 																																																																																																																															
Validaciones																																																																																																																																
Postcondiciones	El sistema debe generar un archivo con formato XLS.																																																																																																																															
Prototipo	<thead> <tr> <th>Ubicación</th> <th>Total Residentes</th> <th>Mayores</th> <th>Menores</th> <th>Blanca</th> <th>Negra</th> <th>Mestiza</th> </tr> </thead> <tbody> <tr><td>Pinar Del Rio</td><td>598.847</td><td>497.736</td><td>101.111</td><td>368.041</td><td>57.023</td><td>36.432</td></tr> <tr><td>Artemisa</td><td>522.215</td><td>438.369</td><td>83.846</td><td>292.846</td><td>49.993</td><td>50.460</td></tr> <tr><td>La Habana</td><td>2.219.533</td><td>1.899.130</td><td>320.403</td><td>1.007.551</td><td>346.500</td><td>378.309</td></tr> <tr><td>Mayabeque</td><td>397.462</td><td>337.778</td><td>59.684</td><td>232.565</td><td>29.982</td><td>36.947</td></tr> <tr><td>Matanzas</td><td>702.889</td><td>594.678</td><td>108.211</td><td>414.752</td><td>73.687</td><td>72.795</td></tr> <tr><td>Villa Clara</td><td>792.199</td><td>670.885</td><td>121.314</td><td>544.184</td><td>42.434</td><td>62.851</td></tr> <tr><td>Cienfuegos</td><td>399.573</td><td>333.882</td><td>65.691</td><td>250.830</td><td>30.615</td><td>46.073</td></tr> <tr><td>Sancti Spiritus</td><td>462.849</td><td>386.587</td><td>76.262</td><td>319.275</td><td>25.171</td><td>29.497</td></tr> <tr><td>Ciego De Avila</td><td>454.990</td><td>382.246</td><td>72.744</td><td>348.917</td><td>29.162</td><td>43.566</td></tr> <tr><td>Camaguey</td><td>834.224</td><td>711.302</td><td>122.922</td><td>479.357</td><td>64.190</td><td>68.832</td></tr> <tr><td>Las Tunas</td><td>553.398</td><td>461.214</td><td>92.184</td><td>342.914</td><td>29.367</td><td>63.260</td></tr> <tr><td>Holguin</td><td>1.099.276</td><td>912.118</td><td>187.158</td><td>663.301</td><td>38.108</td><td>103.300</td></tr> <tr><td>Granma</td><td>885.817</td><td>731.754</td><td>154.063</td><td>243.787</td><td>22.874</td><td>387.273</td></tr> <tr><td>Santiago Cuba</td><td>1.076.400</td><td>886.498</td><td>189.902</td><td>221.817</td><td>199.921</td><td>394.817</td></tr> <tr><td>Guantanamo</td><td>556.268</td><td>453.680</td><td>102.588</td><td>78.400</td><td>81.973</td><td>240.508</td></tr> <tr><td>Isla Juventud</td><td>91.107</td><td>75.438</td><td>15.669</td><td>41.483</td><td>7.776</td><td>19.921</td></tr> <tr><td>Total general</td><td>11.647.047</td><td>9.773.286</td><td>1.873.752</td><td>5.850.020</td><td>1.128.776</td><td>2.034.541</td></tr> </tbody>		Ubicación	Total Residentes	Mayores	Menores	Blanca	Negra	Mestiza	Pinar Del Rio	598.847	497.736	101.111	368.041	57.023	36.432	Artemisa	522.215	438.369	83.846	292.846	49.993	50.460	La Habana	2.219.533	1.899.130	320.403	1.007.551	346.500	378.309	Mayabeque	397.462	337.778	59.684	232.565	29.982	36.947	Matanzas	702.889	594.678	108.211	414.752	73.687	72.795	Villa Clara	792.199	670.885	121.314	544.184	42.434	62.851	Cienfuegos	399.573	333.882	65.691	250.830	30.615	46.073	Sancti Spiritus	462.849	386.587	76.262	319.275	25.171	29.497	Ciego De Avila	454.990	382.246	72.744	348.917	29.162	43.566	Camaguey	834.224	711.302	122.922	479.357	64.190	68.832	Las Tunas	553.398	461.214	92.184	342.914	29.367	63.260	Holguin	1.099.276	912.118	187.158	663.301	38.108	103.300	Granma	885.817	731.754	154.063	243.787	22.874	387.273	Santiago Cuba	1.076.400	886.498	189.902	221.817	199.921	394.817	Guantanamo	556.268	453.680	102.588	78.400	81.973	240.508	Isla Juventud	91.107	75.438	15.669	41.483	7.776	19.921	Total general	11.647.047	9.773.286	1.873.752	5.850.020	1.128.776	2.034.541
Ubicación	Total Residentes	Mayores	Menores	Blanca	Negra	Mestiza																																																																																																																										
Pinar Del Rio	598.847	497.736	101.111	368.041	57.023	36.432																																																																																																																										
Artemisa	522.215	438.369	83.846	292.846	49.993	50.460																																																																																																																										
La Habana	2.219.533	1.899.130	320.403	1.007.551	346.500	378.309																																																																																																																										
Mayabeque	397.462	337.778	59.684	232.565	29.982	36.947																																																																																																																										
Matanzas	702.889	594.678	108.211	414.752	73.687	72.795																																																																																																																										
Villa Clara	792.199	670.885	121.314	544.184	42.434	62.851																																																																																																																										
Cienfuegos	399.573	333.882	65.691	250.830	30.615	46.073																																																																																																																										
Sancti Spiritus	462.849	386.587	76.262	319.275	25.171	29.497																																																																																																																										
Ciego De Avila	454.990	382.246	72.744	348.917	29.162	43.566																																																																																																																										
Camaguey	834.224	711.302	122.922	479.357	64.190	68.832																																																																																																																										
Las Tunas	553.398	461.214	92.184	342.914	29.367	63.260																																																																																																																										
Holguin	1.099.276	912.118	187.158	663.301	38.108	103.300																																																																																																																										
Granma	885.817	731.754	154.063	243.787	22.874	387.273																																																																																																																										
Santiago Cuba	1.076.400	886.498	189.902	221.817	199.921	394.817																																																																																																																										
Guantanamo	556.268	453.680	102.588	78.400	81.973	240.508																																																																																																																										
Isla Juventud	91.107	75.438	15.669	41.483	7.776	19.921																																																																																																																										
Total general	11.647.047	9.773.286	1.873.752	5.850.020	1.128.776	2.034.541																																																																																																																										

Anexo 2. Modelo de dominio del DGM

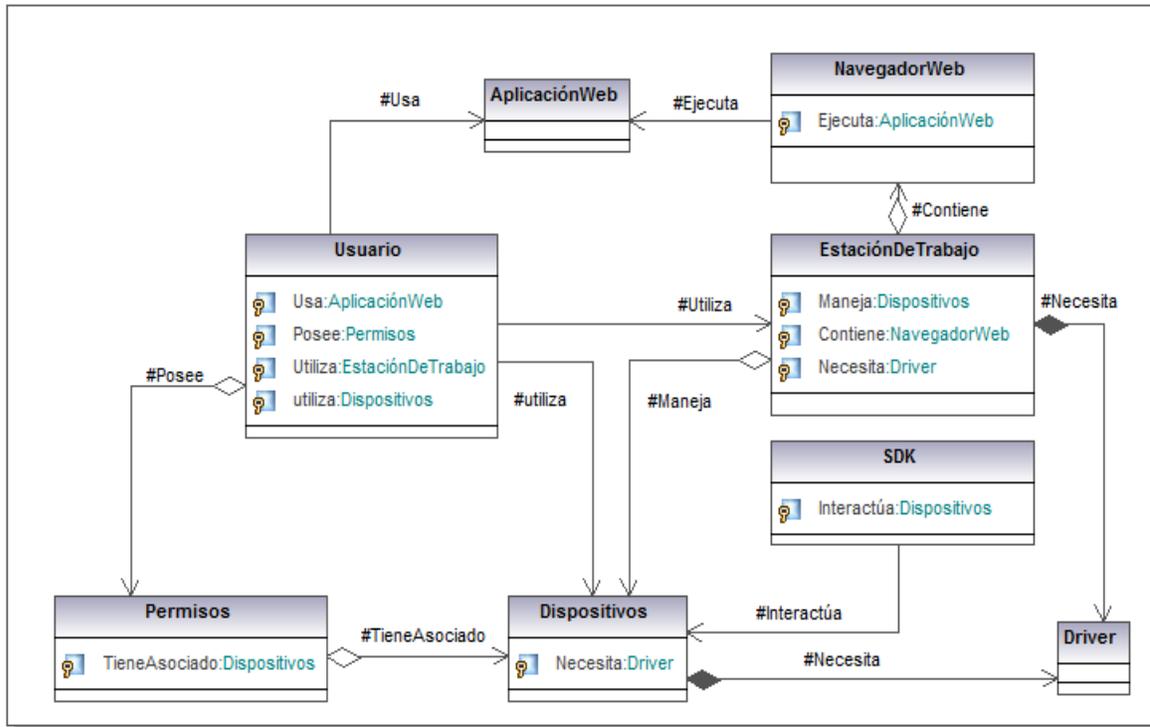


Figura 23. Modelo de dominio DGM.

Anexo 3. Modelos de datos del *data warehouse*

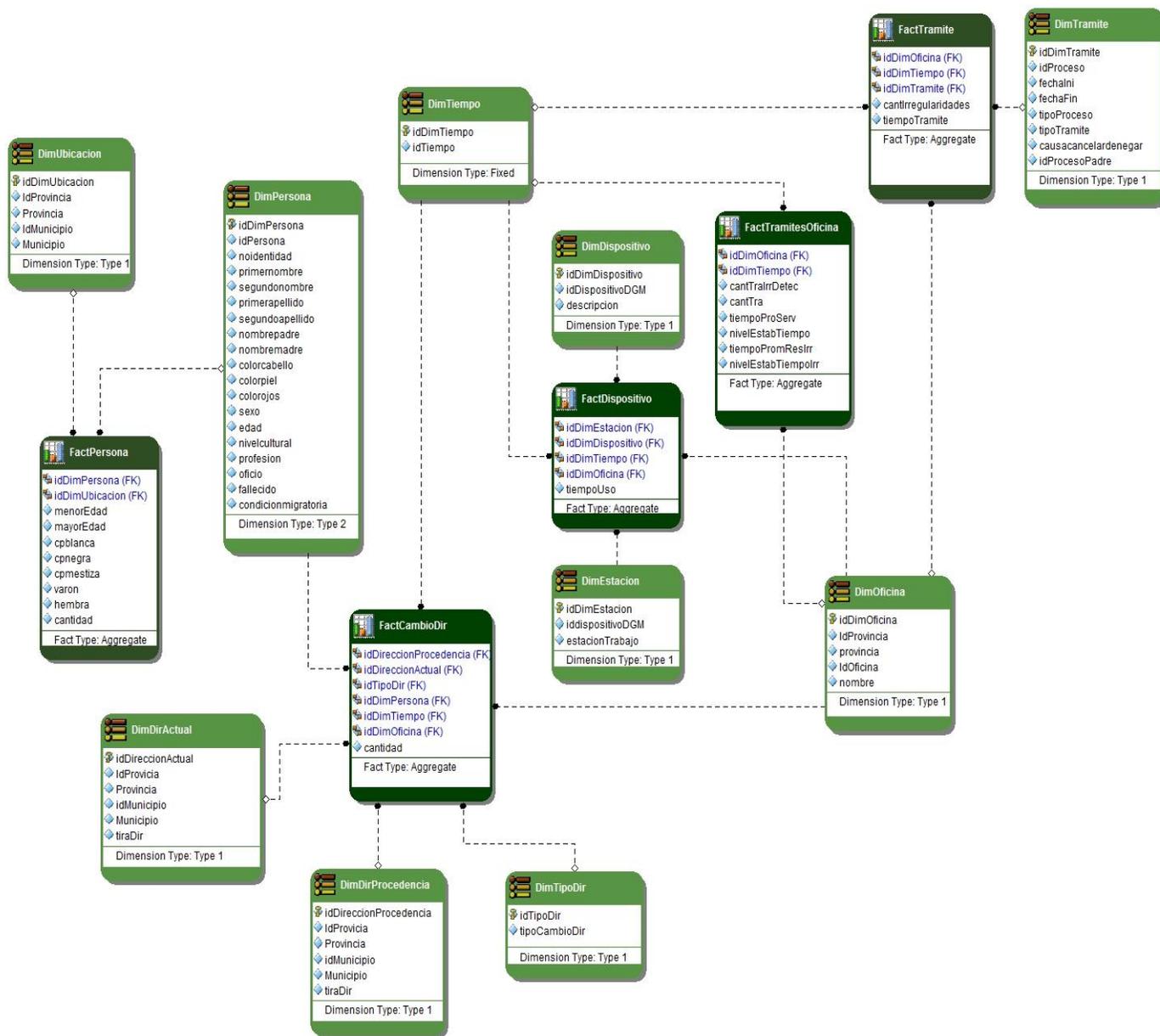


Figura 24. Modelo de datos del *data warehouse*.

Anexo 4. ETL's realizadas

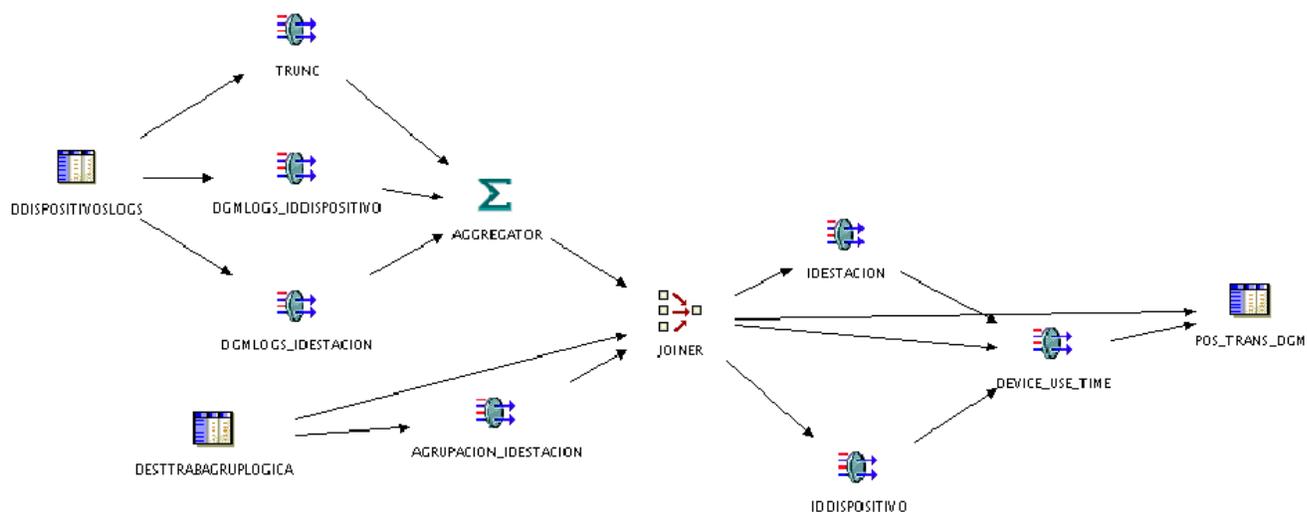


Figura 25. ETL para la carga de los datos del DGM.

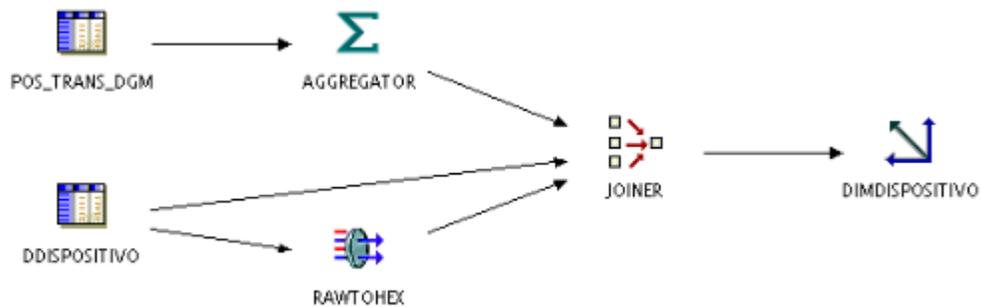


Figura 26. ETL para la carga de los datos hacia DimDispositivo.

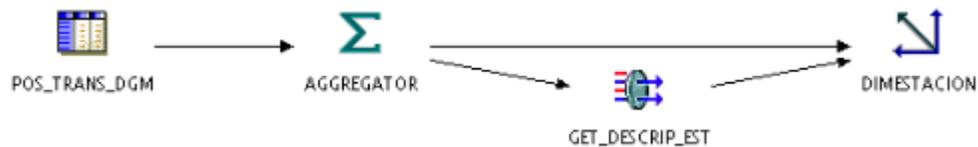


Figura 27. ETL para la carga de los datos hacia DimEstacion.



Figura 28. ETL para la carga de los datos hacia DimOficina.



Figura 29. ETL para la carga de los datos hacia FactDispositivo.

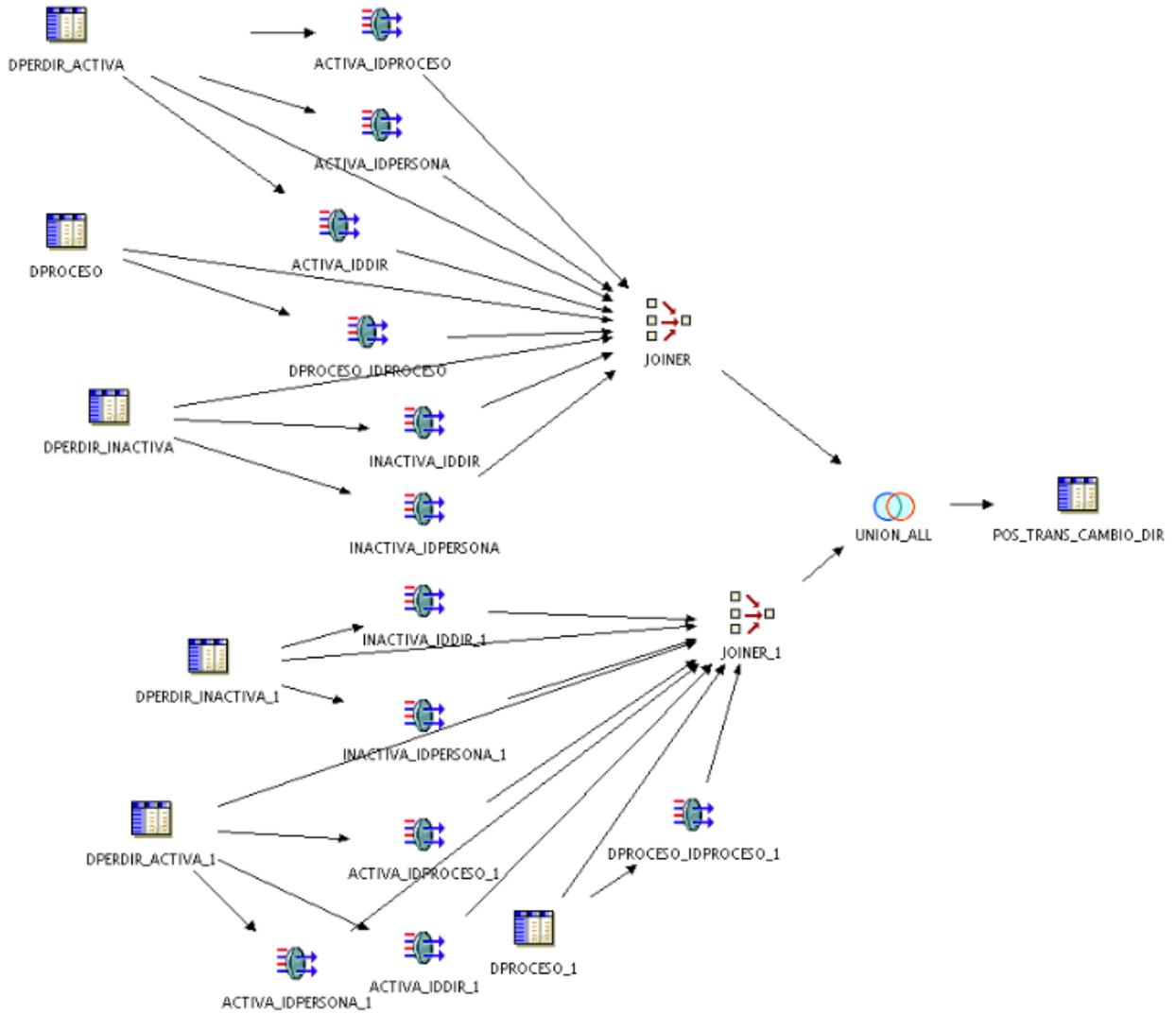


Figura 30. ETL para la realización de la carga de los datos referentes a los cambios de dirección.

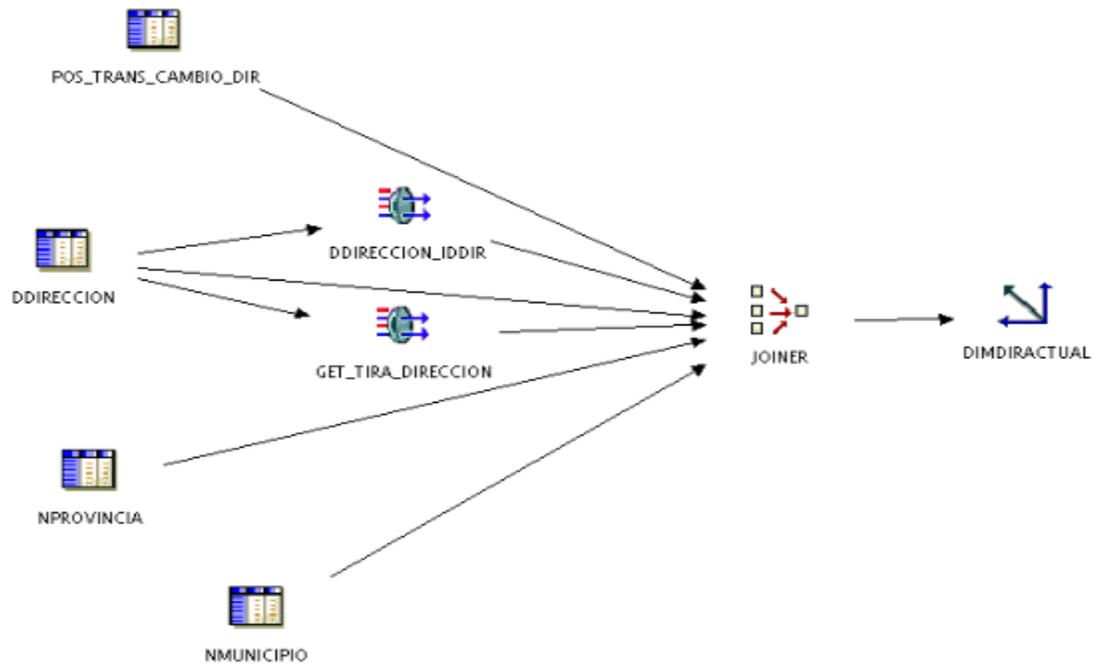


Figura 31. ETL que muestra la carga de los datos hacia la dimensión DimDirActual.

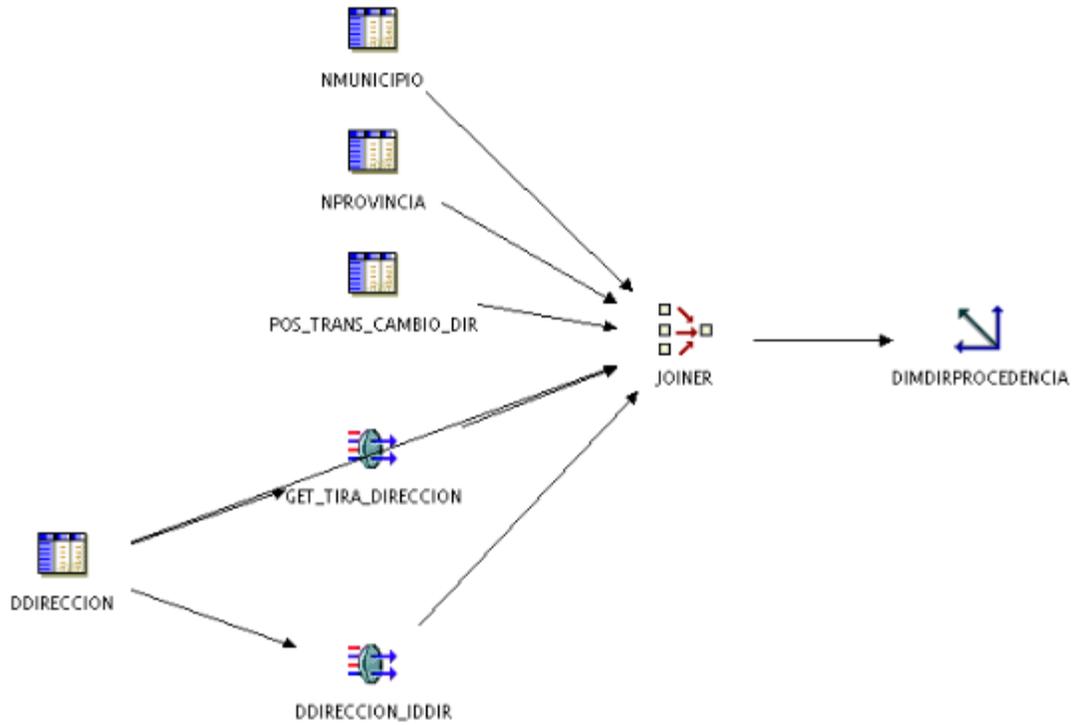


Figura 32. ETL que muestra la carga de los datos hacia la dimensión DimDirProcedencia.

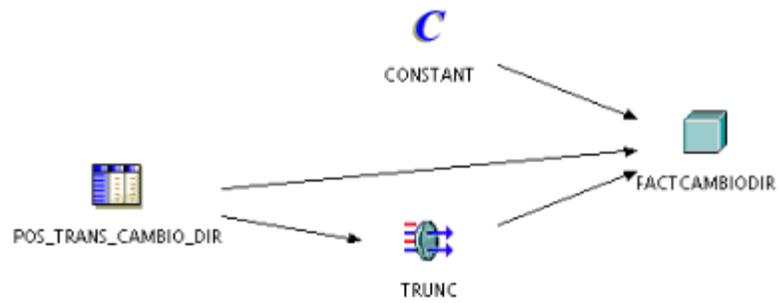


Figura 33. ETL que muestra la carga de los datos hacia la tabla hecho FactCambioDir.

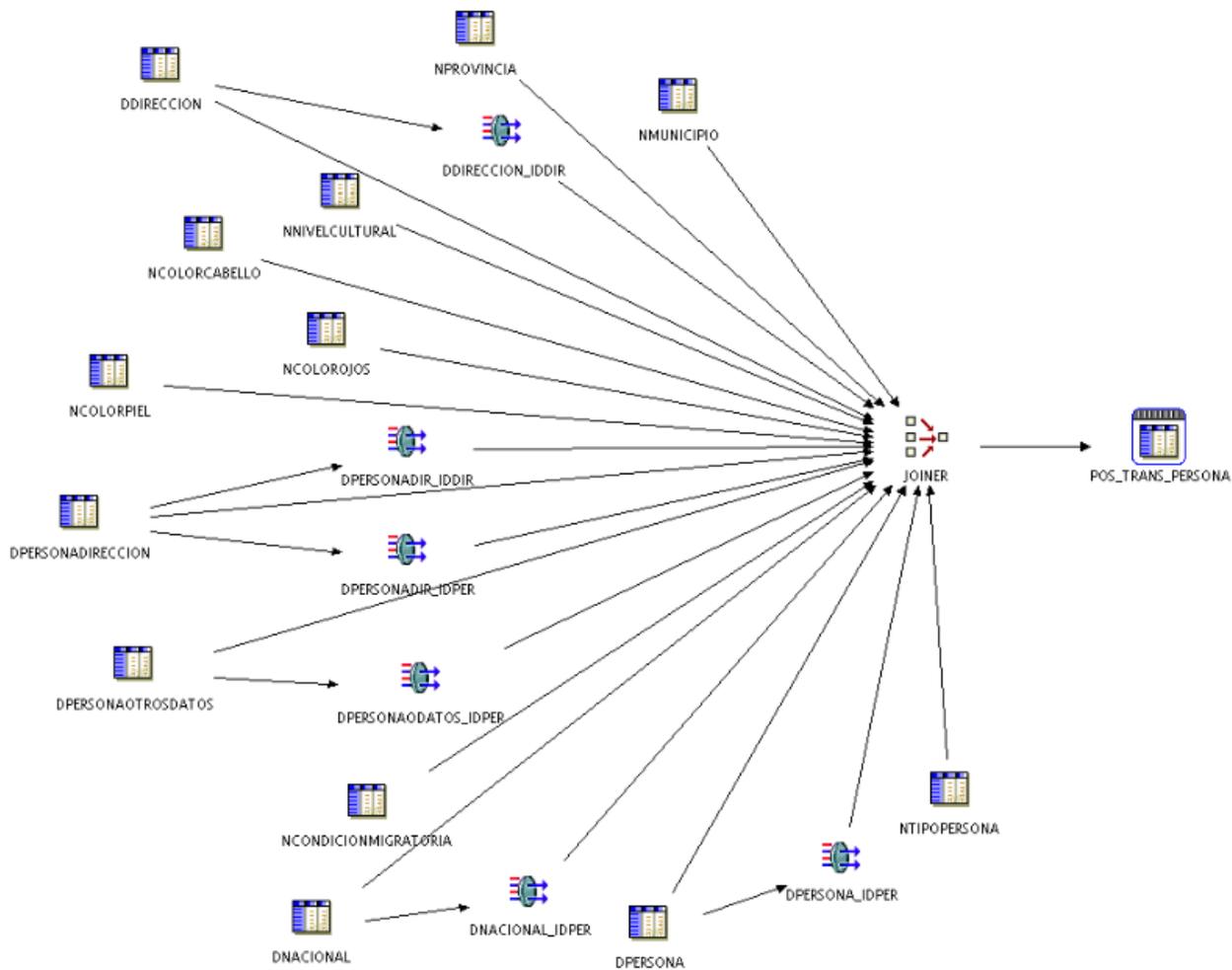


Figura 34. ETL que realiza la carga de los datos referentes a las persona.

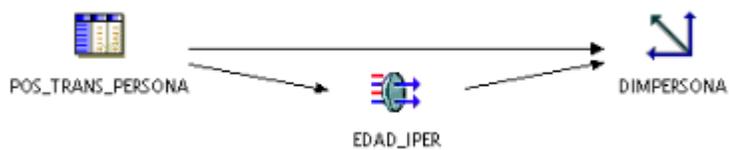


Figura 35. ETL que muestra la carga de los datos hacia la dimensión DimPersona.



Figura 36. ETL que muestra la carga de los datos hacia la dimensión DimUbicación.

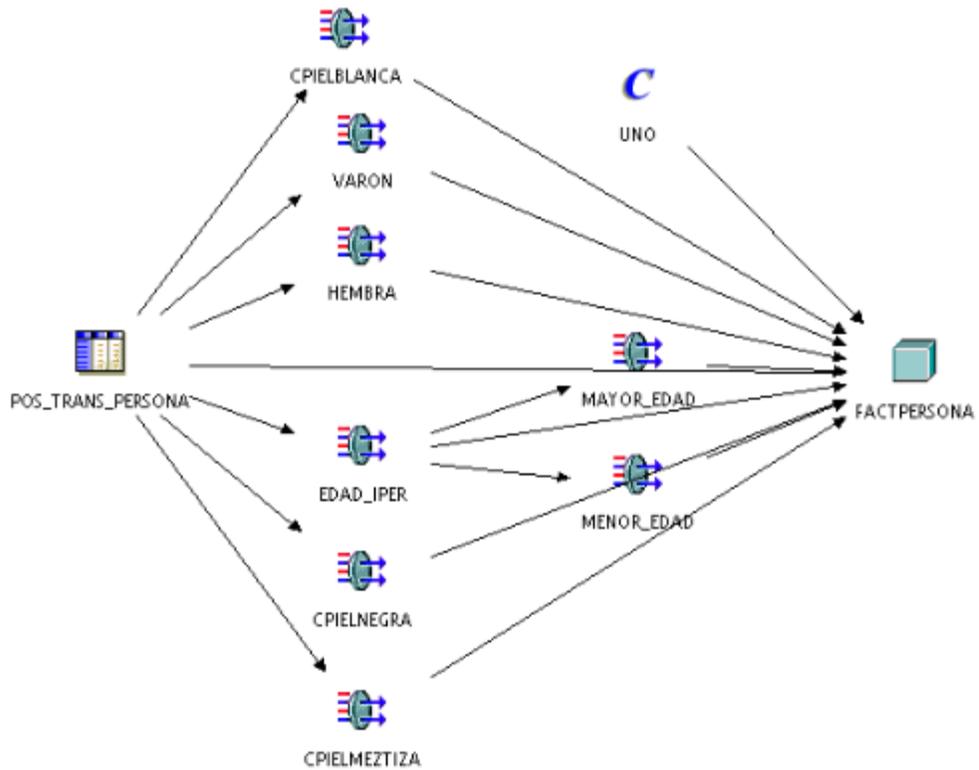


Figura 37. ETL que muestra la carga de los datos hacia la tabla hecho FactPersona.

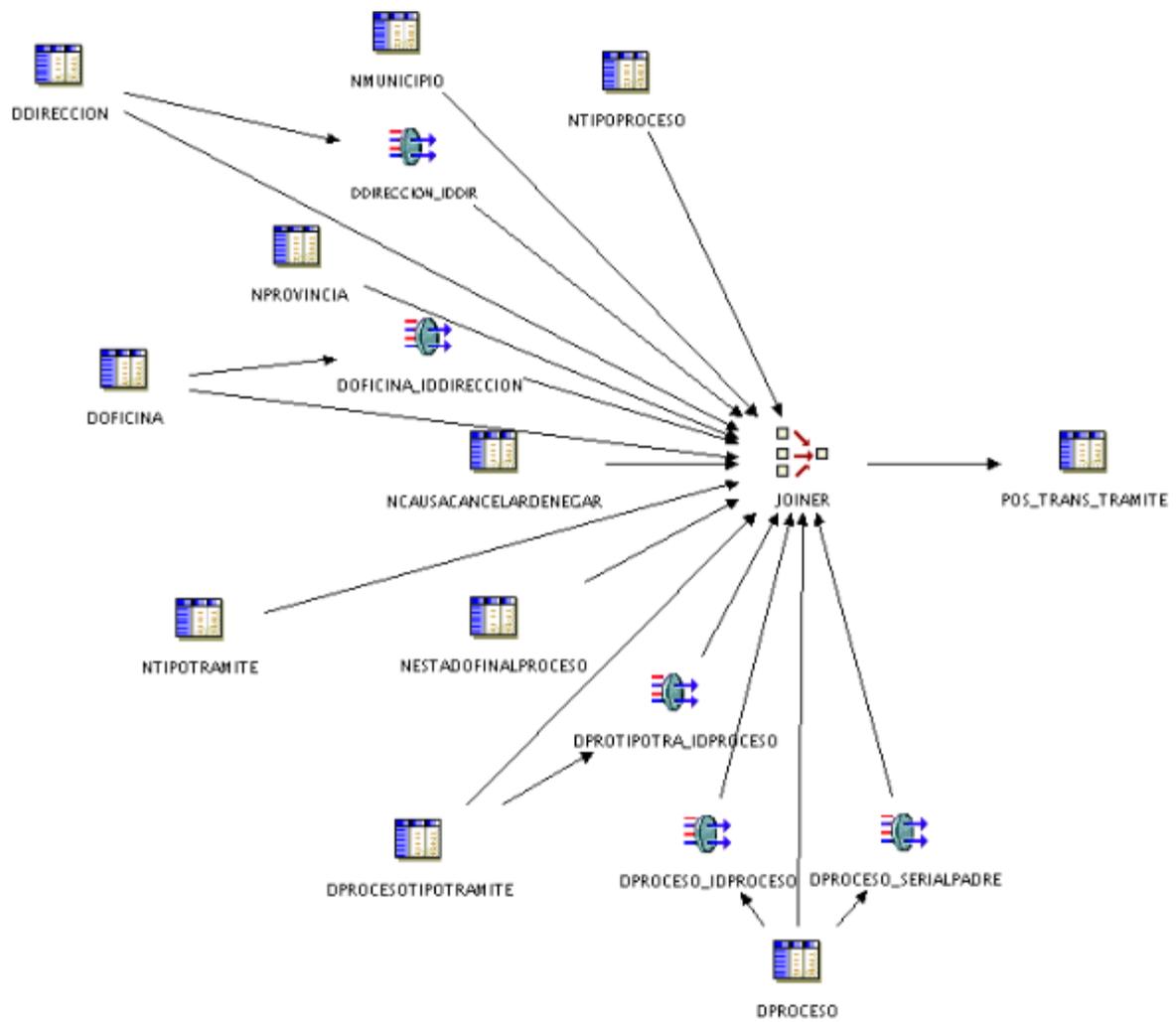


Figura 38. ETL que realiza la carga de los datos referentes a los trámites.

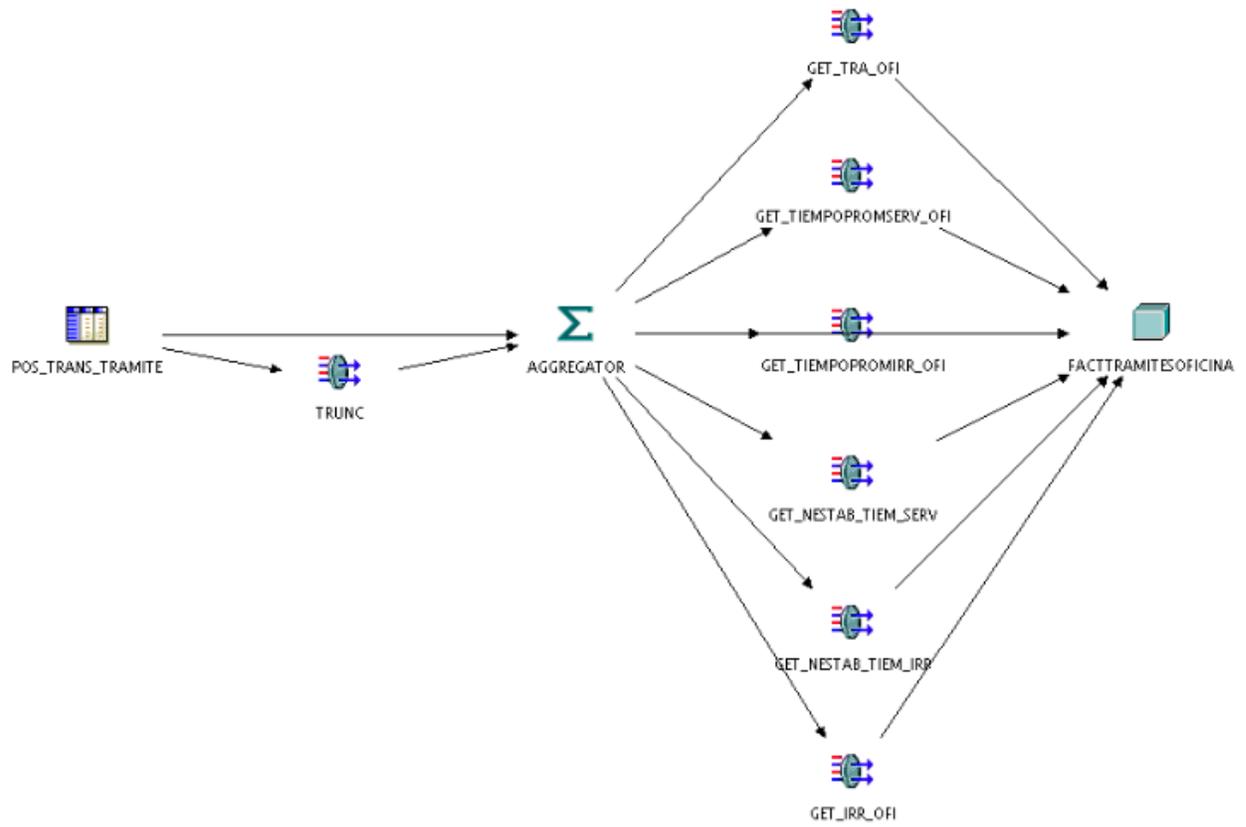


Figura 39.ETL que muestra la carga de los datos hacia la tabla hecho FactTramitesOficina

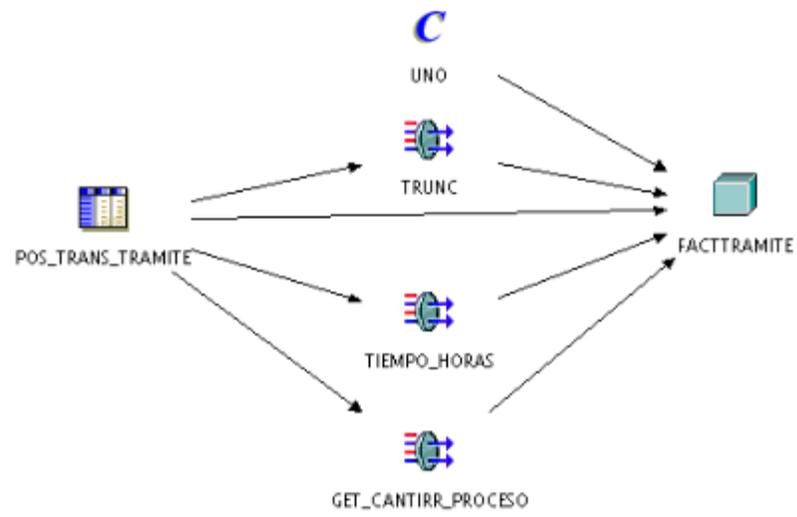
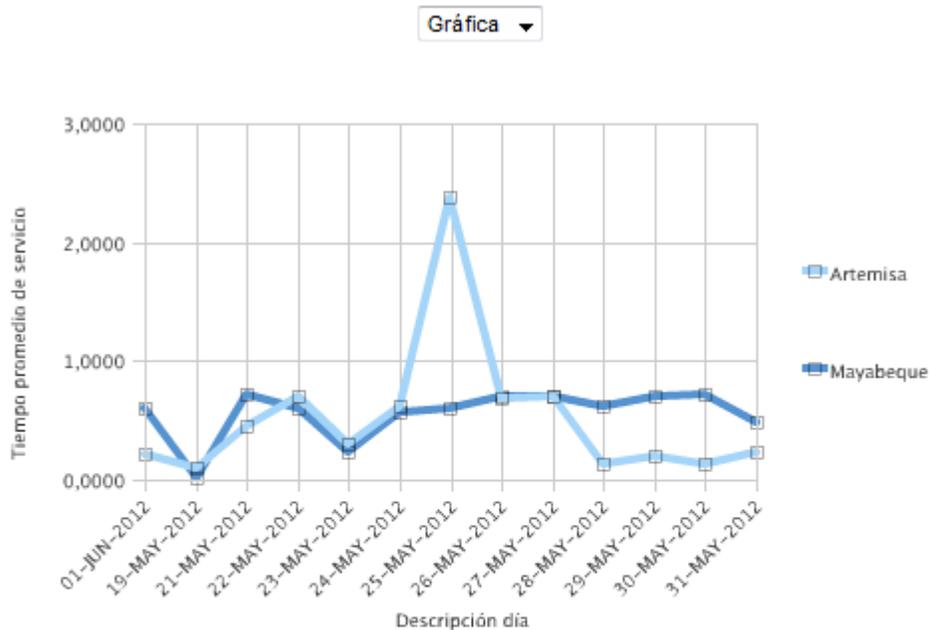


Figura 40. ETL que muestra la carga de los datos hacia la tabla hecho FactTramite.

Anexo 5. Reportes implementados.

Histórico tiempo promedio de servicio



Tiempo promedio de servicio

Provincia	Descripción trimestre	Tiempo promedio de servicio
Artemisa	Trimestre2 2012	0,5279
Mayabeque	Trimestre2 2012	0,5614

Figura 41. Reporte referente al tiempo promedio de servicio.

Coefficiente de Fiabilidad

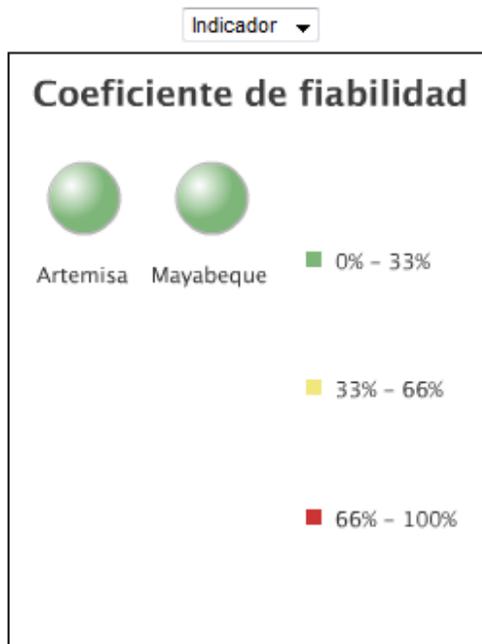


Figura 42. Reporte referente al coeficiente de fiabilidad.

Total de Trámites Realizados por Solicitud de Documentos Actualización

Oficina	Tiempo	Total	Actualización				Solicitud de Documentos			
			Sub Total	Subsanación de Error	Deterioro	Cambio Legal	Sub Total	Salida FAR	Salida Prisiones	Reasentados
Artemisa	Trimestre2 2012	115	31	23	5	3	84	70	14	
Mayabeque	Trimestre2 2012	7	6	4	2		1		1	
Total general		122	37	27	7	3	85	70	15	

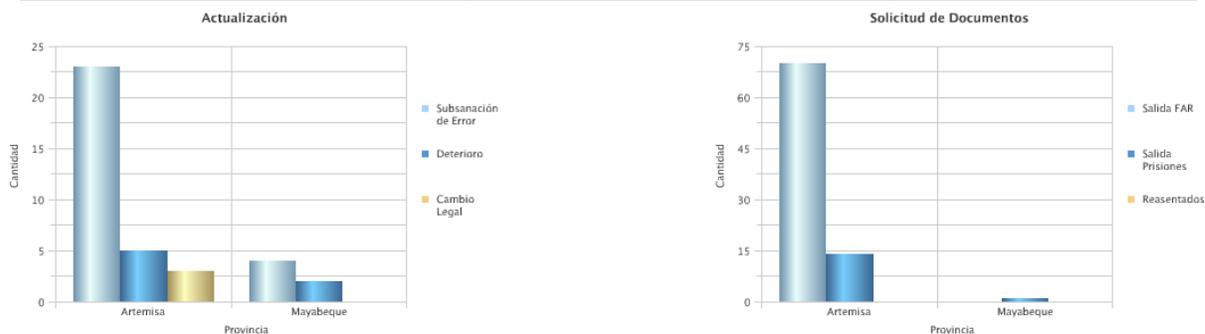


Figura 43. Reporte referente a los trámites de actualización y solicitud de documentos.

Cantidad de Trámites por tipo

Provincia	Total	Tipo de Trámite													
		Actualización de Tarjeta Menor	Arribo	Cambio de dirección	Cambio legal	Confección de pasaporte	Defunción	Desmovilización FAR	Deterioro	Entrada país	Nacimiento CIRP	Nacimiento Hospital	Pérdida	Salida prisión	Subsanación de error
Artemisa	4.034	17	3	300	40	20	112	259	87	9	9	324	29	13	123
Mayabeque	1.408	6	7	79	8	52		100	5	4		29	12	11	35
Total general	5.442	23	10	379	48	72	112	359	92	13	9	353	41	24	158

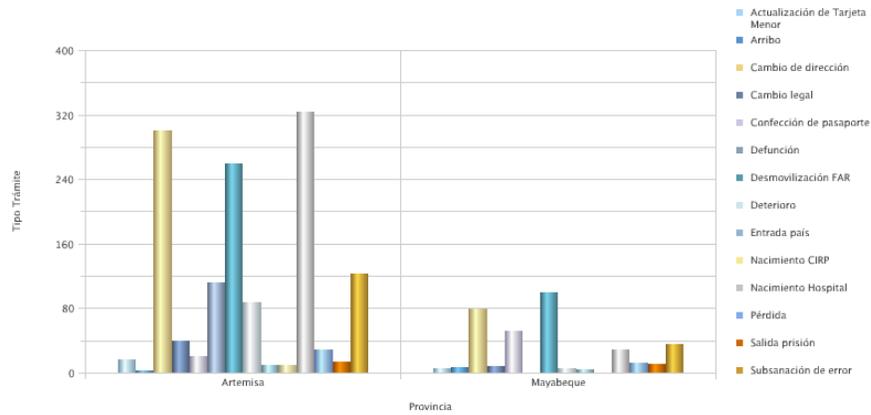


Figura 44. Reporte de trámites por tipo.