



Universidad de las Ciencias Informáticas

Facultad 3

“Desarrollo del módulo de Inteligencia de Negocio en un sistema de control de la producción para una planta de generación de energía distribuida”

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

Autor: Carlos Javier Quintero Blanco
Universidad de las Ciencias Informáticas
(UCI). La Habana Cuba.

Tutores: Ing. Daniel Varona Cordero
Universidad de las Ciencias Informáticas
(UCI). La Habana Cuba.

La Habana, 2013.

Declaración de Autoría

Declaramos ser autores de la presente tesis y reconocemos a la Universidad de las Ciencias Informáticas los derechos patrimoniales de la misma, con carácter exclusivo.

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

Firma del Autor

Firma del Tutor

Agradecimiento

A toda mi familia porque son mi mayor tesoro y sostén. A mi mamá por estar pendiente de cada paso que doy, por su cariño incondicional, su apoyo y dedicación. A mi papá por llenarme de alegría incluso en los peores momentos, por su paciencia, y porque siempre confió en mí. A mis abuelos por hacerme sentir el centro de su mundo, por sus consejos y su amor. A mi hermana por hacerme una mejor persona para tratar de darle un buen ejemplo. A mis tías, tíos, primos y primas que siempre me han dado mucho amor y cariño haciendo de nuestra familia una unión fraternal. A los amigos de toda la vida, a los que están lejos y a los de siempre que nunca fallan en los momentos difíciles. A los que compartieron conmigo su comida, apartamento, alegrías y sentimientos. A Luis Manuel Castillo Cobas por adentrarme desde pequeño en este mundo tecnológico tan ideal para mí. A Daniel Varona Cordero por darme la mano en el momento exacto. A los profesores que marcaron cada etapa de mi aprendizaje. A todos ellos porque la realización de este trabajo se habría quedado en un sueño sin su ayuda. En fin, a todos los que contribuyeron a mi formación como profesional, a los que hicieron que el camino fuera más fácil y siempre confiaron en que podía llegar.

Dedicatoria

A mis padres.

A mis abuelos.

A mis tías y tíos.

A mis primos y primas.

En fin, a toda mi familia, amigas y amigos que siempre confiaron que podría lograrlo y me ayudaron a conseguirlo.

Resumen

Este trabajo consiste en el desarrollo del módulo de Inteligencia de Negocio (GEDIMES BI) para un Sistema de Ejecución de la Manufactura (MES), considerado el corazón de la solución de integración de los sistemas que intervienen en la gestión y control de los procesos de las plantas de generación de energía distribuida en Cuba. Para ello, se desarrolló un modelo de control inteligente para el análisis multidimensional y el monitoreo de la producción, utilizando técnicas de Procesamiento Analítico en Línea (OLAP). Además, se diseñó un Cubo OLAP basado en los hechos críticos del proceso de producción de la planta, en el que cada métrica de las tablas de hechos son completadas a partir de un servicio de integración que actúa sobre el universo de los Sistemas Transaccionales (OLTP), permitiendo obtener los indicadores claves de rendimiento (KPI) que muestran de forma gráfica el estado del proceso productivo para ayudar en la toma de decisiones ante las dispersiones ocurridas.

PALABRAS CLAVES

Indicadores, OLAP, Modelo multidimensional

Índice de Contenido

Introducción	1
1 Capítulo 1. Fundamentación Teórica.	5
1.1 Generación de Energía.....	5
1.1.1 Métodos de generación	5
1.1.2 Flujo tecnológico.....	6
1.1.3 Sistema de control básico de una planta de Generación de Energía Distribuida... ..	7
1.1.4 Generación de energía distribuida en La Habana	7
1.2 Inteligencia de Negocio.....	8
1.2.1 Sistemas de soporte a la Inteligencia de Negocio.....	8
1.2.2 Norma ISO 62264.....	9
1.3 Almacén de datos	14
1.3.1 Necesidad de los almacenes de datos.....	14
1.3.2 Procesamiento Transaccional en Línea	15
1.3.3 Procesamiento Analítico en Línea.....	15
1.3.4 ETL.....	18
1.3.5 ¿Cómo funciona todo?.....	¡Error! Marcador no definido.
1.4 Herramientas y Metodologías	19
1.4.1 SQL Server.....	20
1.4.2 Lenguaje de consulta MDX.....	21
1.4.3 Metodología.....	21
1.5 Conclusiones	23
2 Capítulo 2. Propuesta de solución	25
2.1 Descripción de la Infraestructura en la capa superior de Control	25
2.2 GEDIMES.....	27
2.3 Requisitos.....	28
2.3.1 Requisitos del negocio.....	28
2.3.2 Restricciones del negocio	29
2.4 Modelado Dimensional (Cubo MOLAP)	31
2.5 Diseño Físico del Almacén de Datos.....	33
2.5.1 Diseño físico de los hechos	33

2.5.2	Diseño físico de las dimensiones	39
2.6	Indicadores Claves de Rendimiento.....	44
2.7	Extracción, Transformación y Carga (ETL)	48
2.8	Implementación	50
2.8.1	Microsoft Excel 2010.....	50
2.8.2	Microsoft SQL Server Agent (Agente SQL Server)	53
2.9	Infraestructura de la solución de Inteligencia de Negocio.....	54
2.10	Conclusiones	55
3	Capítulo 3. Evaluación del análisis de datos	56
3.1	Validación y pruebas.....	56
3.2	Resultados de la aplicación de las pruebas	57
3.2.1	Resultados de la prueba de integridad de los datos.....	57
3.2.2	Resultados de la prueba de carga y estrés	58
3.2.3	Resultados de la prueba del modelo	59
3.2.4	Validación del Sistema.....	59
3.3	Conclusiones	60
	Conclusiones	61
	Recomendaciones	62
	Anexos.....	65

Índice de Figuras

Figura 1. Flujo tecnológico de una planta de generación de energía distribuida.....	6
Figura 2. Diseño de los procesos tecnológicos.	7
Figura 3. Jerarquía funcional de la Norma ISO 62264.....	10
Figura 4. Modelo del Manejo de las Operaciones de Manufactura.....	11
Figura 5. Modelo de las actividades para el Manejo de las Operaciones de Manufactura.....	12
Figura 6. Categorías de intercambio de la información	13
Figura 7. Modelo de jerarquía física.....	14
Figura 8. Ejemplo de un Cubo OLAP.....	16
Figura 9. Integración de un sistema OLAP.....	19
Figura 10. Tareas de la Metodología de Kimball	23
Figura 11. Diagrama de Clases del Sistema GEDIMES.....	28
Figura 12. Arquitectura física de hardware para implementar el sistema.....	30
Figura 13. Diseño Multidimensional del Cubo MOLAP de GEDIMES.....	33
Figura 14. Relación multidimensional del hecho Producción.....	34
Figura 15. Relación multidimensional del hecho Mantenimiento	36
Figura 16. Relación multidimensional del hecho Laboratorio.....	37
Figura 17. Relación multidimensional del hecho Inventario.....	38
Figura 18. Jerarquía de la Dimensión Equipo	39
Figura 19. Jerarquía de la Dimensión RRHH	40
Figura 20. Jerarquía de la Dimensión MateriaPieza.....	41
Figura 21. Jerarquía de la Dimensión Prueba	42
Figura 22. Jerarquía de la Dimensión OrdenTrabajo.....	43
Figura 23. Jerarquía de la Dimensión Factura	44
Figura 24. Arquitectura del Servicio de Integración.....	49
Figura 25. Diseño del flujo de integración para la dimensión RRHH.	49
Figura 26. Tabla Pívor Multidimensional usando Microsoft Excel 2010	51
Figura 27. Gráfico Multidimensional en forma de columna usando Microsoft Excel 2010.....	52
Figura 28. Gráfico Multidimensional en forma de línea usando Microsoft Excel 2010	52
Figura 29. Gráfico Multidimensional en forma de pastel usando Microsoft Excel 2010.....	53
Figura 30. Configuración del Agente SQL Server para ejecutar un flujo de integración.....	54
Figura 31. Distribución genérica de la red para la Solución BI en GEDICH.....	54
Figura 32. Por ciento de errores por iteración	57

Figura 33. Datos que se encuentran persistidos en la base de datos.....	58
Figura 34. Datos consultados desde el cliente OLAP.....	58
Figura 35. Muestra el tiempo medio de respuesta por la cantidad de usuarios	59
Figura 36. Diseño del flujo de integración para la dimensión RRHH.	65
Figura 37. Diseño del flujo de integración para la dimensión Factura.....	65
Figura 38. Diseño del flujo de integración para la dimensión Prueba.	66
Figura 39. Diseño del flujo de integración para el hecho Producción.	66
Figura 40. Acta de aceptación de la solución Gedimes BI.....	67

Índice de Tablas

Tabla 1. Relación de Servidores en la Plataforma GEDIMES.	26
Tabla 2. Relación de OLTP que alimenta el Cubo MOLAP GEDIMES y su Aplicación Consola GEDIMES BI	27
Tabla 3. Operaciones basadas en la implementación de los requisitos.....	29
Tabla 4. Matriz de procesos/dimensiones (Bus Matrix)	32
Tabla 5. Hecho Producción.....	34
Tabla 6. Hecho Mantenimiento.	36
Tabla 7. Hecho Laboratorio.....	37
Tabla 8. Hecho Inventario.	38
Tabla 9. Dimensión Equipo.	39
Tabla 10. Dimensión RRHH.....	40
Tabla 11. Dimensión MateriaPieza.....	41
Tabla 12. Dimensión Prueba.....	41
Tabla 13. Dimensión OrdenTrabajo	42
Tabla 14. Dimensión Factura.	43
Tabla 15. Indicadores Claves de Rendimiento.....	48

Introducción

A partir del año 2005 la dirección del estado cubano decidió dar solución al deterioro que sufría el Sistema Eléctrico Nacional (SEN). Este estaba centralizado en una decena de plantas de generación térmica cuyo sistema tecnológico provenía del ya extinto campo socialista. Se decidió pasar a un sistema distribuido de plantas cuyas características minimizaban el consumo de combustibles, también el costo de dichas plantas era muy inferior al de una planta de generación térmica y además los proveedores internacionales de esta tecnología suministraban un alto nivel de automatización.

Por lo cual el cambio de la infraestructura tecnológica de generación para cubrir la demanda energética del pueblo cubano trajo consigo un aumento en la diversidad de plantas geográficamente distribuidas a lo largo de todo el territorio nacional, produciéndose un aumento en la complejidad de la gestión de los procesos que acompañan la explotación y mantenimiento de estas tecnologías, tanto a nivel de planta como a nivel de provincia y de país.

Para enfrentar dicha complejidad desde el punto de vista metodológico, la dirección de generación distribuida de la Unión Eléctrica (UNE) diseñó e implantó un Manual de Gestión (MAGEST) (1) para el gobierno de estos procesos.

Las plantas de generación de energía distribuida poseen un precario sistema tecnológico para la gestión de la información en la toma de decisiones, basada metodológicamente en el MAGEST, sin embargo cuentan con un alto nivel de automatización a nivel del control básico de la planta soportado por un sistema Siemens que culmina su esquema de control de procesos tecnológicos en dos estaciones de operaciones y una de ingeniería, ubicadas en la Sala de Control Remoto (RCMS) de dichas plantas, que a su vez tienen un sistema de Control de Supervisión y Adquisición de Datos (SCADA) que le permite al operador la visualización, monitoreo y control de las operaciones básicas de la planta.

Este esquema posee las siguientes características:

- 1- Es un sistema estable altamente distribuido por una red Profinet de autómatas SIMATIC 315 2DN\DP que gobiernan los elementos fundamentales de dichas plantas.
- 2- Posee un alto nivel de variables de actuación y monitoreo asociados a un conjunto elevado de alarmas de procesos que permiten al operador mantener la estabilidad del proceso tecnológico.

Introducción

3- La elevada complejidad del tratamiento de la información a nivel de SCADA, no permite al personal de la planta en ningún momento determinar la eficiencia del proceso. Es por ello que acuden a soluciones de ofimática para realizar el tratamiento de los históricos y así poder obtener indicadores que den ideas del comportamiento global del sistema y les permita entonces tomar decisiones.

4- Los subsistemas tecnológicos de gestión de los procesos empresariales que existen en este momento (Mantenimiento, Recursos Humanos, Economía, etc.) poseen una naturaleza semiestática y aislada, ya que en ningún momento están integrados a la dinámica de las variables del proceso productivo.

Se hace necesario para los expertos del dominio y los grupos que toman decisiones de esta organización, convertir la información en un recurso estratégico de la entidad, cuya gestión y respaldo, conviertan a la empresa en un sistema adaptable, que va aprendiendo, y evoluciona a partir de sus propios conocimientos, los cuales son resultados del tratamiento de los recursos de información almacenados en los registros de las interacciones de cada proceso o subsistema que conforma el sistema en su totalidad.

A partir del análisis de esta problemática se identifica el siguiente **problema a resolver**: El actual proceso de toma de decisiones para la generación de energía distribuida en la planta de Ciudad de la Habana presenta dificultades asociadas al consumo de información desagregada que son soportadas por distintos sistemas especializados.

La investigación se orienta en el **objeto de estudio**: Sistemas de Inteligencia de Negocio aplicados a Sistemas de Generación de Energía Distribuida, planteándose el **objetivo general**: Desarrollar un módulo de inteligencia de negocio para un sistema de control de la producción en una planta de generación de energía distribuida. El **campo de acción** se ubica en la toma de decisiones en la generación de energía distribuida en la planta de Ciudad de la Habana.

Idea a defender:

Si se desarrolla un módulo de Inteligencia de Negocio constituido por un Cubo de Procesamiento Analítico en Línea (OLAP) que por medio de una herramienta de Extracción, Transformación y Carga (ETL) integre los parámetros de monitoreo suministrados por los distintos sistemas especializados de la planta de Generación de Energía Distribuida Ciudad de la Habana (GEDICH) se facilitará el proceso de toma de decisiones asociado a la generación de energía distribuida en la mencionada planta.

Introducción

Para garantizar el cumplimiento del objetivo se plantean las siguientes **tareas** de investigación y desarrollo:

1. Definición del marco teórico investigativo y estado del arte.
2. Diseño de la propuesta de solución.
3. Desarrollo del Almacén de Datos que responda a la gestión integrada de los procesos que intervienen en una planta de generación de energía distribuida utilizando la norma ISO/IEC 62264.
4. Desarrollo del Cubo de Procesamiento Analítico en Línea que responda al proceso de toma de decisiones para la generación de energía distribuida.
5. Implementación del mecanismo de Integración, utilizando tecnología ETL.
6. Verificación de la propuesta de solución.
7. Validación del modelo utilizado para el desarrollo de la aplicación de Inteligencia de Negocio.

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

Métodos teóricos:

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 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 la planta de generación de energía distribuida GEDICH.
- Brindar un documento con la especificación de requisitos del sistema desarrollado.

Introducción

- Manual de Usuario de la aplicación.

El presente trabajo se estructura en tres capítulos descritos a continuación:

En el **Capítulo 1 Fundamentación Teórica** se definen y analizan los conceptos o elementos relacionados con los sistemas de inteligencia organizacional posibilitando una mejor comprensión de las características de los mismos y una valoración de las posibles vías de solución a la problemática planteada. Se realiza una descripción del estudio y selección de las herramientas a utilizar en el desarrollo de la propuesta de solución, así como la metodología que soporta la construcción de un almacén de datos, que guiará todo el proceso.

En el **Capítulo 2 Propuesta de solución** se abordan todos los elementos de la propuesta desde la concepción de la solución, la descripción de los requisitos del negocio y las restricciones del negocio, el modelado dimensional del cubo OLAP, el diseño físico del almacén de datos y la implementación de la solución.

En el **Capítulo 3 Evaluación del análisis de datos** se muestra el análisis de los resultados de la validación de la solución de inteligencia de negocios para GEDICH. 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.

1 Capítulo 1. Fundamentación Teórica.

En el presente capítulo se definen y analizan los conceptos o elementos relacionados con los sistemas de inteligencia organizacional posibilitando una mejor comprensión de las características de los mismos y una valoración de las posibles vías de solución a la problemática planteada. Se realiza una descripción de las herramientas a utilizar en el desarrollo de la propuesta de solución, así como la metodología que soporta la construcción de un almacén de datos, que guiará todo el proceso.

1.1 Generación de Energía

La generación de energía eléctrica es una actividad humana básica, pues está directamente relacionada con los requerimientos actuales del hombre. Todas las formas de utilización de las fuentes de energía, tanto las habituales como las denominadas alternativas o no convencionales, agreden, en mayor o menor medida, el ambiente, siendo de todos modos la generación de energía eléctrica una de las actividades humanas que causan menor impacto.

En general, la generación de energía eléctrica consiste en transformar alguna clase de energía química, mecánica, térmica o luminosa, entre otras, en energía eléctrica. Para lograr la generación industrial se recurre a instalaciones denominadas centrales eléctricas, que ejecutan alguna de las transformaciones citadas. Estas constituyen el primer escalón del sistema de suministro eléctrico.

Dependiendo de la fuente primaria de energía utilizada, las centrales generadoras se clasifican en termoeléctricas (de carbón, petróleo, gas, nucleares y solares termoeléctricas), hidroeléctricas (aprovechando las corrientes de los ríos o del mar: mareomotrices), eólicas y solares fotovoltaicas. La mayor parte de la energía eléctrica generada a nivel mundial proviene de los dos primeros tipos de centrales reseñados (2).

1.1.1 Métodos de generación

La generación de energía puede ser centralizada o distribuida. Esta última consiste, básicamente, en la generación de energía eléctrica por medio de muchas pequeñas fuentes de energía. La definición más global de la generación distribuida específica que es aquella que se conecta a la red de distribución de energía eléctrica y que se caracteriza por encontrarse instalada en puntos cercanos al consumo.

Capítulo 1. Fundamentación Teórica.

La generación distribuida aparece como alternativa al modelo centralizado tradicional, el cual presenta una fuerte dependencia energética. A través de sistemas con tecnologías cada vez más eficientes energéticamente hablando y de la incorporación de fuentes renovables de producción de energía, entra en juego la generación distribuida, desplazando los centros de generación de energía a los puntos de consumo de ésta. (3)

Este sistema de generación aporta una serie de beneficios respecto al modelo centralizado destacando la reducción de costos inherentes al transporte de la energía, a los que también hay que sumar la mayor eficiencia energética de los equipos de generación, como los nuevos equipos compactos de cogeneración que presentan un rendimiento muy superior a los equipos de las plantas convencionales ya que la antigüedad de estos equipos se puede establecer en una media de entre 20 y 50 años.

A todo esto hay que añadir que dichos equipos están cada vez más automatizados, necesitándose así una menor intervención humana para su manejo y monitorización. Este hecho permite tener un mayor control de la gestión de la demanda favoreciendo el autoconsumo y el balance neto.

1.1.2 Flujo tecnológico

En la figura siguiente se muestra un esquema simplificado del flujo tecnológico principal que se lleva a cabo en la central eléctrica generadora de energía distribuida.

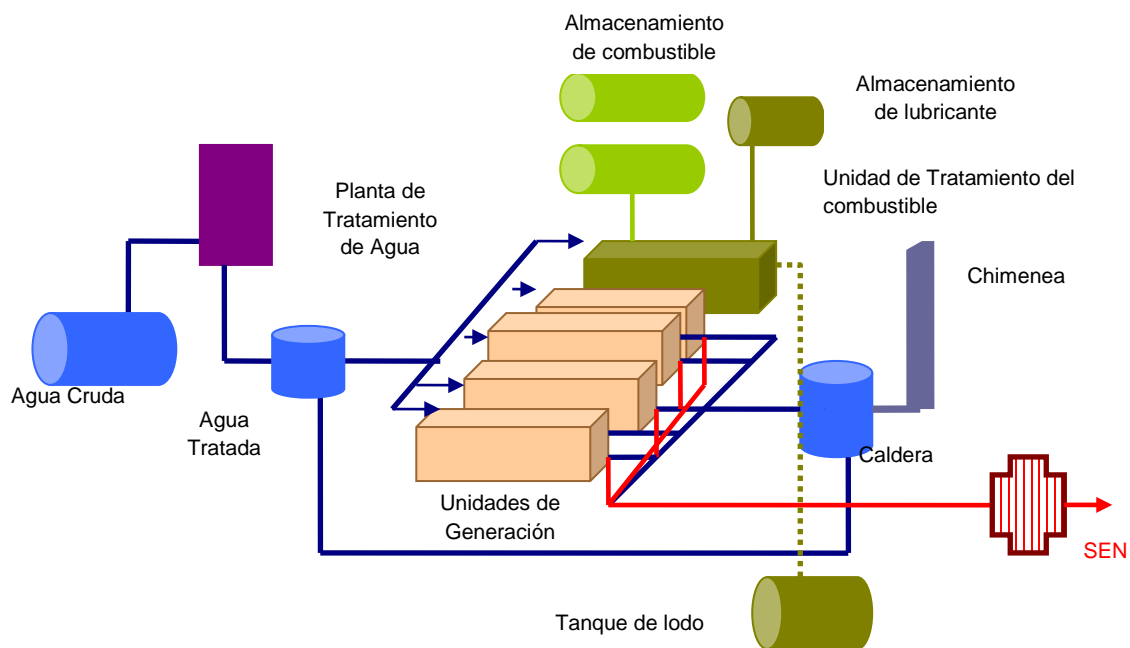


Figura 1. Flujo tecnológico de una planta de generación de energía distribuida.

Capítulo 1. Fundamentación Teórica.

Estas centrales poseen una Unidad de control que responde a un sistema de automatización distribuido, con autómatas producidos por la empresa SIEMENS, los cuales, en cada caso, son una solución degenerada de *Las Soluciones de Automatización Homogénea con SIMATIC*, con un nivel de redundancia, ya sea por seguridad o disponibilidad, según requiera la situación.

La pretensión final de este tipo de diseño está asociada al control de los procesos tecnológicos mediante las Tecnologías de la Información y las Comunicaciones (TIC) a nivel empresarial, como lo demuestra la siguiente imagen extraída de los manuales tecnológicos de la SIEMENS.

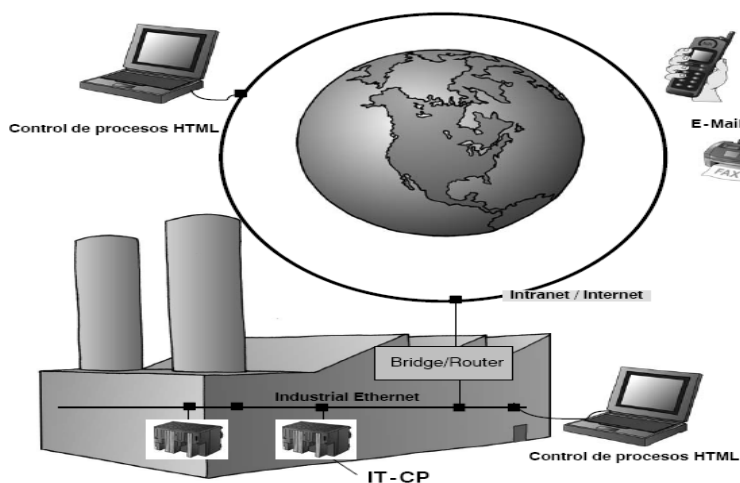


Figura 2. Diseño de los procesos tecnológicos

1.1.3 Sistema de control básico de una planta de Generación de Energía Distribuida

Las plantas de producción de energía distribuida en Cuba, con régimen base, están basadas fundamentalmente en motores de combustión interna. Forman una distribución del equipamiento por baterías. Poseen una lógica de control distribuida a nivel de planta a través de una red Profinet de autómatas 315 2DN/DP de SIEMENS para controlar los procesos inherentes a cada motor generador y otros elementos fundamentales a nivel del piso de producción. En esta red también están conectadas 3 estaciones de operación en un centro de control remoto (RCMS), que regentan el proceso tecnológico a nivel de planta por medio de un Sistema de Control de Supervisión y Adquisición de Datos (SCADA), ejecutando una estación SIMATIC WinCC 6.2 sobre plataforma Windows XP Sp2.

1.1.4 Generación de energía distribuida en La Habana

En La Habana, en Abril del 2007, surgió la Empresa de Generación Eléctrica (GEDICH), con el objetivo de gestionar estas tecnologías, proporcionando así un sistema de dimensión provincial

Capítulo 1. Fundamentación Teórica.

que fuera capaz de responder ante el SEN como una unidad de gestión técnica administrativa de todas las instalaciones interrelacionadas con este fin en la provincia. Esta empresa, en perfeccionamiento empresarial, se rige por el decreto ley No. 281 del Comité Ejecutivo del Consejo de Ministros del 7 de Agosto de 2007. Su objetivo es implantar un Sistema de Gestión Empresarial basado en la utilización de las TIC.

1.2 Inteligencia de Negocio

La Inteligencia de Negocio tiene como objetivo facilitar la toma de decisiones y en el momento correcto. Para ello gestiona el conocimiento. Es soportada por tecnologías de la información que incluyen herramientas de soporte a la toma de decisiones y la contribución de expertos que comparan y analizan tendencias utilizando los indicadores que en ese momento se requieren (4). Estas herramientas colaboran con la Inteligencia de Negocios, ya que asisten el análisis y la presentación de los datos.

La Inteligencia de Negocio, según la Norma ISO 62264 para la Gestión Empresarial (5), es soportada por los sistemas que se describen a continuación.

1.2.1 Sistemas de soporte a la Inteligencia de Negocio

1.2.1.1 Sistemas de Información Gerencial

Los Sistemas de Información Gerencial (en inglés Manufacturing Information System, MIS) se diferencian de los sistemas de información comunes en que, para analizar la información, usan otros sistemas que utilizan las actividades operacionales de la organización, la comprensión del funcionamiento actual de la empresa, para ofrecer conocimientos y así respaldar las decisiones empresariales.

1.2.1.2 Sistemas de Planificación de Recursos Empresariales

Otras de las herramientas usadas para el control de las actividades empresariales son los Sistemas de Planificación de Recursos Empresariales (ERP), estos integran y manejan muchos de los negocios asociados con las operaciones de producción y de los aspectos de distribución de una empresa en la producción de bienes y servicio. Los sistemas ERP pueden intervenir en el control de las actividades como la producción, mantenimientos, administración de inventarios, pruebas de calidad y la administración de recursos humanos, que forman parte de las actividades críticas identificadas en la generación de energía.

1.2.1.3 Sistemas de Control de Supervisión y Adquisición de Datos

Capítulo 1. Fundamentación Teórica.

En el caso de los sistemas para el control de la producción es necesario que exista un alto nivel de automatización como complemento al proceso. Para ello son utilizados los Sistemas de Control de Supervisión y Adquisición de Datos (SCADA). Estos son basados en computadores que permiten supervisar y controlar variables del proceso a distancia, proporcionando comunicación con los dispositivos de campo (controladores autónomos) y controlando el proceso de forma automática por medio de un software especializado. También proveen toda la información que se genera en el proceso productivo a diversos usuarios, tanto del mismo nivel como de otros usuarios supervisores dentro de la empresa (supervisión, control de la calidad, control de la producción, almacenamiento de datos, etc.).

1.2.1.4 Sistema de Ejecución de Manufactura

Para lograr que los sistemas antes descritos puedan ser integrados, se debe crear un Sistema de Ejecución de Manufactura (MES) que dirija y monitoree los procesos de producción en la planta, incluyendo el trabajo manual o automático, así como preguntas on-line y enlaces a las tareas que tienen lugar en la planta de producción. El MES es un intermediario entre el sistema ERP y el SCADA. El módulo de inteligencia de negocio es el centro de la solución MES.

1.2.2 Norma ISO 62264

Al aplicar normas ISO en la institución, unido a una estrategia de comunicación interna, se pueden determinar en cada subsistema las variables de entrada y salida de cada subproceso, así como las variables y los informes a tener en cuenta para medir el estado real de la planta.

Este proceso normalizado es la Norma ISO/IEC 62264 (5), que define las interfaces entre las actividades empresariales y el control de la producción. Su objetivo es operar los sistemas de gestión de la información con una fácil integración. En la siguiente figura se muestra la interfaz antes descrita. Esta frontera es identificada usando modelos relevantes que representan funciones, equipamiento físico, e información dentro del dominio para el Manejo de las Operaciones de Manufactura (MOM).

Capítulo 1. Fundamentación Teórica.

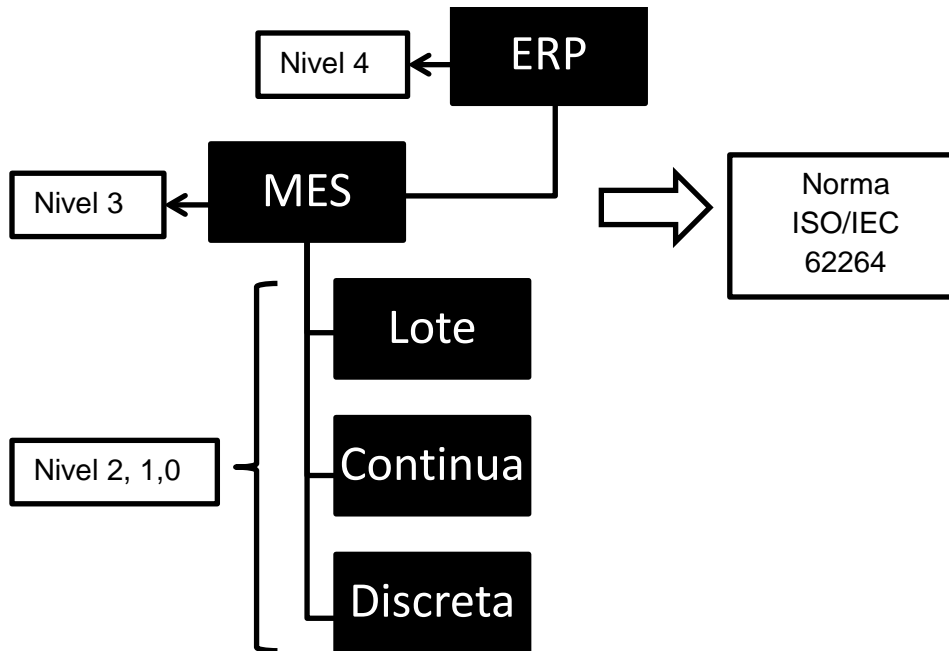


Figura 3. Jerarquía funcional de la Norma ISO 62264.

El nivel 0 pertenece al nivel físico donde se encuentra el equipamiento tecnológico. En el caso de la generación eléctrica la producción es continua, ya que el producto no se almacena ni se cuantifica por unidades. En el nivel 1 se encuentran los sensores y actuadores usados para monitorear el proceso. En el nivel 2 se indica el control de las actividades de forma manual o automática, manteniendo el proceso estable o bajo control. En el nivel 3 se analizan los datos, se registran los mantenimientos, inventarios y los procesos de calidad para poder generar recetas e indicadores que reflejen el estado actual de la planta, lo cual permite el monitoreo y control del proceso de producción. Por último, el nivel 4 establece el plan de producción básico de la planta y los recursos usados en la producción.

En las áreas sombreadas en gris de la siguiente figura se representan las actividades MOM antes mencionadas. Las MOM son una colección de manejos sobre las Operaciones de Producción, Mantenimiento, Calidad e Inventario, así como otras actividades relacionadas con la producción.

Capítulo 1. Fundamentación Teórica.

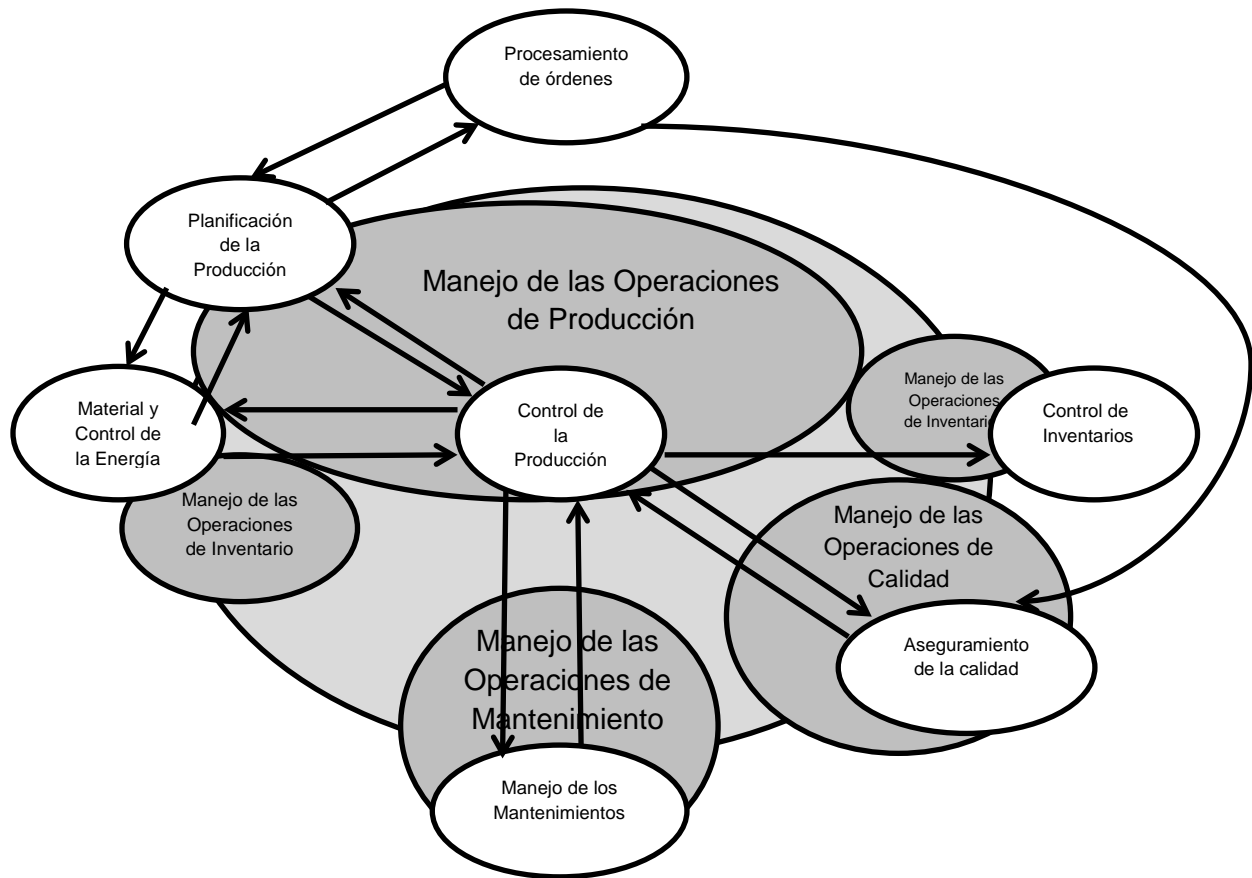


Figura 4. Modelo del Manejo de las Operaciones de Manufactura

Los modelos para el manejo de las operaciones de producción incluyen las actividades del control de la producción, así como la planificación de la producción que operan a nivel 3, a este mismo nivel se encuentran el manejo de las operaciones de los mantenimientos y calidad. Por último, para los inventarios se tienen en cuenta las actividades del manejo de inventarios y materiales, incluyendo el control de los productos inventariados y el control de las actividades de materias y energía, definidas como operaciones de este mismo nivel. (5)

Las actividades MOM coordinan el personal, equipamiento, material y energía para producir un producto, con el costo, cantidad, calidad y tiempo requerido. Los flujos de trabajo que soportan las MOM están representados entre el nivel 3 y 4.

Para definir los detalles de las actividades en todas las categorías de las aplicaciones MOM, sus interrelaciones dentro del nivel 3 y con las aplicaciones del nivel 4, se muestra la siguiente figura, que expone los modelos definidos y usados para elaborar las cuatro (4) actividades críticas del negocio: **Producción, Mantenimiento, Pruebas de Calidad, y Manejo de**

Capítulo 1. Fundamentación Teórica.

Inventarios. Las funciones y el alto nivel de intercambios son definidos para cada actividad MOM.

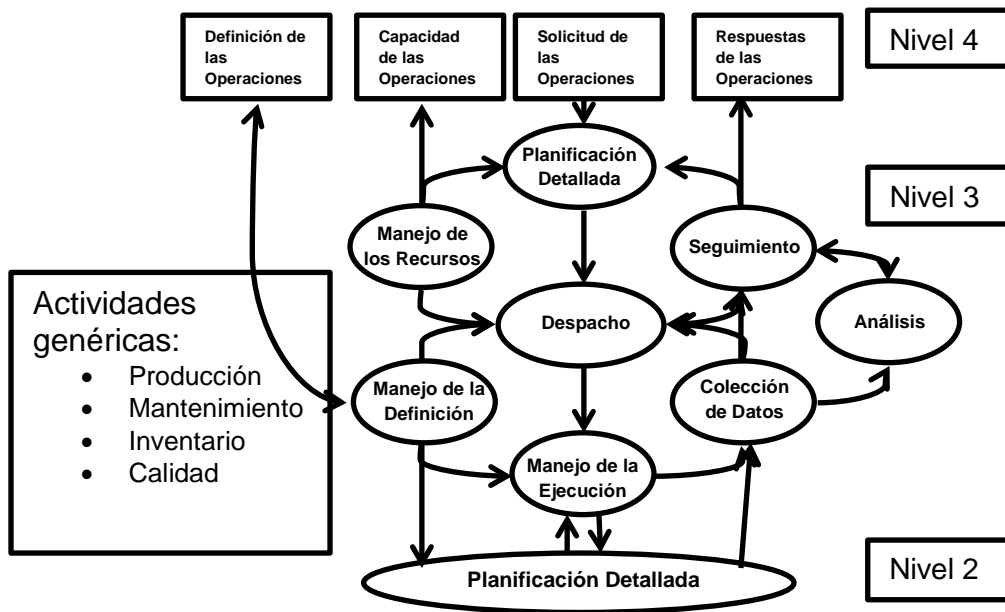


Figura 5. Modelo de las actividades para el Manejo de las Operaciones de Manufactura

El modelo de actividades genéricas define un ciclo general de solicitudes y respuestas que empiezan con las solicitudes o planificaciones, convirtiéndolas entonces en planificación detallada y trabajo de despacho, manejando la ejecución del trabajo y colectando datos los cuales retornan convertidos en las respuestas. Estos ciclos son soportados por el análisis del trabajo realizado, el manejo de los recursos usados en la ejecución de la producción y el manejo de las definiciones del trabajo realizado. (5)

El propósito de estos modelos es identificar los posibles flujos de datos dentro de las operaciones de producción. Los óvalos en el modelo indican la colección de tareas identificadas como las principales actividades y las líneas con flechas representan la importancia de los flujos de información entre las actividades.

La información de la producción contiene cuatro categorías que deben intercambiar entre el sistema de planificación del negocio y el sistema de control de las operaciones de producción. En la siguiente Figura se muestra dicho modelo de objetos.

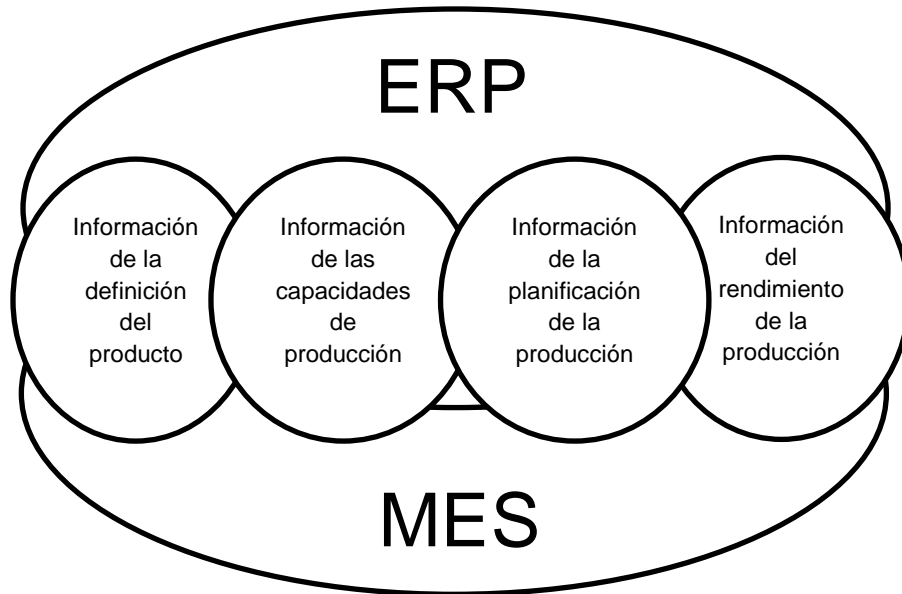


Figura 6. Categorías de intercambio de la información

La información de definición del producto hace referencia a qué se definirá para hacer el producto. La información de las capacidades de producción detalla cuáles son los recursos disponibles. La información de la planificación de la producción especifica cuál es la producción actual que será ejecutada y la información del rendimiento de la producción especifica cuál fue la producción lograda. Estas categorías se aplican a la información de las operaciones de producción y control de la información. (5)

La estructura de los niveles de dominio funcional contiene la jerarquía de la aplicación dentro de la empresa. En la siguiente Figura se muestran los niveles de la jerarquía física asociados con los equipamientos, que incluyen elementos de los procedimientos de trabajo y los recursos asociados requeridos.

Capítulo 1. Fundamentación Teórica.

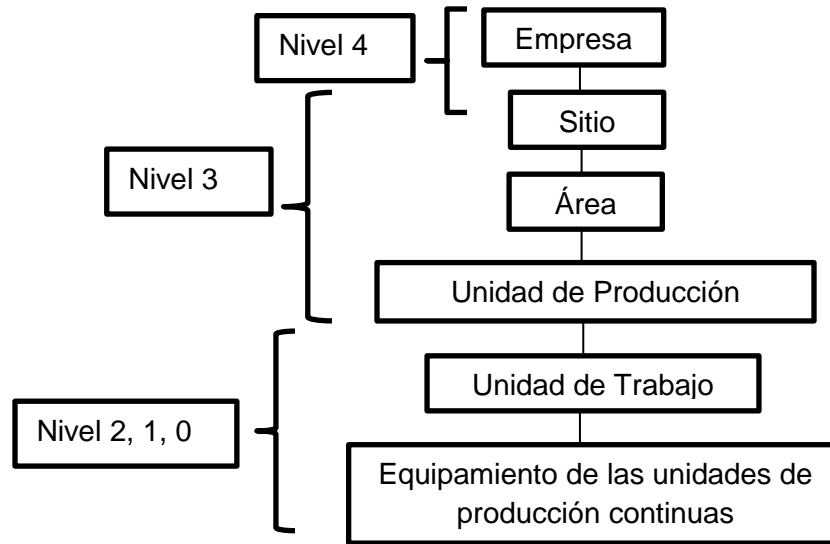


Figura 7. Modelo de jerarquía física

La empresa es una colección de uno o más sitios o áreas, que tiene la responsabilidad de determinar qué productos van a ser producidos. Las funciones de nivel 4 son referentes a la empresa y sus sitios, aunque para la planificación se pueden utilizar otras áreas representadas en el nivel 3. Los sitios constituyen un grupo físico, geográfico o lógico determinado por la empresa, que contiene áreas y sus respectivas jerarquías (5). Estos generalmente definen las capacidades de producción. Las áreas son el espacio donde ocurren las actividades más importantes del nivel 3. Estas realizan los niveles más bajos que soportan las operaciones de producción.

1.3 Almacén de datos

Un almacén de datos (Data Warehouse, DW) es una colección de datos orientada a un determinado ámbito (empresa, organización, etc.), integrado, no volátil y variable en el tiempo, que ayuda a la toma de decisiones en la entidad en la que se utiliza. Se trata, sobre todo, de un historial completo de la organización, más allá de la información transaccional y operacional, almacenada en una base de datos diseñada para favorecer el análisis y la divulgación eficiente de datos (especialmente con herramientas OLAP, de procesamiento analítico en línea). (6)

1.3.1 Necesidad de los almacenes de datos

La proliferación de sistemas de información sustentados en bases de datos ha generalizado el uso de herramientas que permiten obtener informes complejos, resúmenes e incluso estadísticas globales sobre la información almacenada con el objetivo de asistir en la toma de

Capítulo 1. Fundamentación Teórica.

decisiones. Por tanto, cada día es más necesario distinguir dos usos diferentes del sistema de información: el procesamiento transaccional (OLTP) y el procesamiento analítico (OLAP).

1.3.2 Procesamiento Transaccional en Línea

El procesamiento transaccional en tiempo real (por su término en inglés, On-Line Transactional Processing) constituye el trabajo primario en un sistema de información. Este trabajo consiste en realizar transacciones, es decir, actualizaciones y consultas a la base de datos con un objetivo operacional: hacer funcionar las aplicaciones de la organización, proporcionar información sobre el estado del sistema de información y permitir actualizarlo conforme va variando la realidad del contexto de la organización. Muestras de este tipo de trabajo transaccional son, por ejemplo, en el caso de una empresa, la inserción de un nuevo cliente, el cambio de sueldo de un empleado, la tramitación de un pedido, el almacenamiento de una venta, la impresión de una factura, la baja de un producto, etc. Es el trabajo diario y para el que inicialmente se ha diseñado la base de datos. (7)

1.3.3 Procesamiento Analítico en Línea

Conocido por su término en inglés, On-line Analytical Processing es una solución que se utiliza en la inteligencia de negocio para agilizar las consultas de grandes cantidades de datos y que emplea estructuras multidimensionales denominadas Cubos OLAP con datos resumidos de grandes Bases de Datos o Sistemas OLTP. Un aspecto distintivo de OLAP es que las consultas tienen gran velocidad de respuesta. Realiza un procesamiento analítico en tiempo real, engloba un conjunto de operaciones, exclusivamente de consultas, en las que se requiere agregar y cruzar gran cantidad de información. El objetivo de estas consultas es realizar informes y resúmenes, generalmente para el apoyo en la toma de decisiones. Ejemplos de este tipo de trabajo analítico pueden ser resúmenes de ventas mensuales, los consumos eléctricos por días, la espera media de los pacientes en cirugía digestiva de un hospital, el producto cuyas ventas han crecido más en el último trimestre, las llamadas por horas, etc. Este tipo de consultas suelen emanarse de los departamentos de dirección, logística o prospectiva y requieren muchos recursos. (7)

1.3.3.1 Cubo OLAP

Un Cubo OLAP es una base de datos multidimensional, en la cual el almacenamiento físico de los datos se realiza en un vector multidimensional. Se puede considerar como una ampliación de las dos dimensiones de una hoja de cálculo. (6)

Capítulo 1. Fundamentación Teórica.

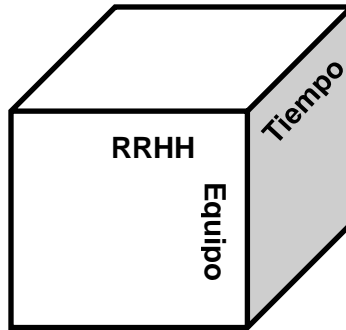


Figura 8. Ejemplo de un Cubo OLAP

Cada dimensión de un Cubo OLAP puede resumirse mediante una jerarquía. Por ejemplo, si se considera una escala (o dimensión) temporal "Mayo de 2013" se puede incluir en "Segundo Trimestre de 2013", que a su vez se incluye en "Año 2013". De igual manera, otra dimensión de un cubo que refleje un recurso humano, los nombres se pueden agrupar por cargos o categoría, los equipos podrían clasificarse por tipos y las partidas de gastos podrían agruparse en tipos de gastos. En cambio, el analista podría comenzar en un nivel muy resumido, como, por ejemplo, el total de la diferencia entre los resultados reales y lo presupuestado, para posteriormente descender en el cubo (en sus jerarquías) para poder observar con un mayor nivel de detalle que le permita descubrir en el cubo los lugares en los que se ha producido esta diferencia, según los productos y períodos.

1.3.3.2 Operadores OLAP

Sobre los modelos multidimensionales se pueden realizar operaciones que permiten a los usuarios explorar e investigar los datos en busca de respuestas; estas son (6):

Roll-up (Generalización Progresiva): Cambiar una categoría en la granularidad por una categoría menos fina, es decir, es ir de lo específico a lo general. Permite apreciar los datos en menor nivel de detalle, subiendo por una jerarquía definida en un cubo. Esto brinda la posibilidad de quitar un nivel o criterio de agregación en el análisis, adicionando los grupos actuales.

Drill-down (Profundización Progresiva): Es lo inverso de Roll-Up, es decir, es ir de lo general a lo específico. Permite apreciar los datos en un mayor detalle, bajando por una jerarquía definida en un cubo. Esto brinda la posibilidad de introducir un nuevo nivel o criterio de agregación en el análisis, disgregando los grupos actuales.

Capítulo 1. Fundamentación Teórica.

Drill-across (Profundización a través de): Cruzar más de una tabla de hechos. Es muy similar al funcionamiento de Drill-down, con la diferencia de que Drill-across no se realiza sobre una jerarquía, sino que agrega un atributo a la consulta como nuevo criterio de análisis.

Roll-across (Generalización a través de): Es muy similar al funcionamiento de Drill-up, con la diferencia de que Roll-across no se hace sobre una jerarquía, sino que elimina un atributo de la consulta, excluyendo de esta manera un criterio de análisis.

Slice (Rebanar): Imponer condiciones sobre las dimensiones.

Pivot (Rotar): Elegir atributos para la tabla de salida y cambiar la disposición de los atributos. Permite seleccionar el orden de visualización de los atributos e indicadores con el objetivo de analizar la información desde diferentes perspectivas.

1.3.3.3 Tipos de Sistemas OLAP

Se pueden definir varios tipos de Sistemas OLAP, dependiendo de las técnicas que se utilicen a la hora de obtener los datos y la forma en la que están estructurados y almacenados. Entre los principales Sistemas OLAP se encuentran:

ROLAP (Procesamiento Analítico Relacional en Línea): Se trata de sistemas y herramientas OLAP construidas sobre una base de datos relacional. La premisa de los sistemas ROLAP es que las capacidades OLAP se soportan mejor contra las bases de datos relacionales. Soporta técnicas de optimización de accesos para acelerar las consultas. Estas optimizaciones son, entre otras, el particionado de los datos a nivel de aplicación, el soporte a la desnormalización, el precalculado de datos e índices y uniones múltiples. (4)

MOLAP (Procesamiento Analítico Multidimensional en Línea): Se trata de una alternativa a la tecnología ROLAP que utiliza un modelo de datos multidimensional. Requiere un preprocesamiento y almacenamiento de la información contenida en el Cubo OLAP, que almacena estos datos en una matriz multidimensional. Un sistema MOLAP usa una base de datos, en la que la información se almacena y visualiza multidimensionalmente. (4)

El tipo de sistema OLAP seleccionado es el MOLAP ya que se tienen distintos sistemas fuentes, lo cual incrementa la complejidad de la gestión del proceso de carga de la información en el almacén de datos.

1.3.3.4 Beneficios de la tecnología OLAP

La tecnología OLAP proporciona un conjunto de beneficios relacionados a continuación (7):

Capítulo 1. Fundamentación Teórica.

Flexibilidad: Con reportes OLAP se tiene al alcance una gran variedad de configuraciones, proporcionándoles a los usuarios un control y flexibilidad sin precedentes. El acceso a los datos y las labores de generación de cubos se presentan ante los analistas como una tarea sencilla. Estos reportes permiten realizar agregaciones y combinaciones de los datos de maneras mucho más complejas y ambiciosas, con objetivos de análisis más estratégicos.

Ágil toma de decisiones: Los reportes OLAP facilitan la toma de decisiones, ya que la generación y modificación de reportes se hace de manera rápida con el Sistema OLAP. La creación y modificación de cubos puede ser realizada por los mismos usuarios finales de manera rápida y ágil. Esta es una gran ventaja para las organizaciones donde los usuarios finales necesitan cubos altamente personalizados. Con los métodos tradicionales de reportes OLAP, los cubos son generados para grandes grupos de personas, obligando a los usuarios a examinar los montones de datos para encontrar la información que estaban buscando. Los reportes OLAP empoderan a los usuarios para que puedan generar cubos más específicos, que suplen sus necesidades de información. Si el usuario final quiere adicionar o retirar algo del cubo, puede hacerlo.

Decisiones inteligentes: Los reportes OLAP ofrecen mayor poder analítico, ya que los cubos de reporte contienen cálculos que van más allá de los simples conteos, promedios y sumas. Para complementar estos cálculos, existe una amplia variedad de gráficos estadísticos y analíticos, que le permiten realizar una mirada más profunda a sus datos, y aumentar el poder visual de los gráficos.

Mayor productividad: Los usuarios finales de la organización, los desarrolladores de los departamentos de tecnologías de la información y, consecuentemente, de toda la organización, pueden experimentar una mayor productividad, un acceso más controlado y oportuno de la información de carácter estratégico, permitiendo realizar un proceso de toma de decisiones más efectivas.

Reducción de la carga de trabajo: En el personal de los departamentos de tecnologías de la información se presentará una reducción de carga de trabajo de desarrollo de aplicaciones, al hacer que los usuarios finales sean lo bastante autosuficientes, como para poder realizar sus propios cambios de esquema y construir sus propios modelos.

1.3.4 ETL

Finalmente, si se ha decidido diseñar un almacén de datos, y ya estando implementado mediante tecnología ROLAP o MOLAP, el siguiente paso es cargar los datos. El proceso

Capítulo 1. Fundamentación Teórica.

tradicional de base de datos más parecido a la carga de un almacén de datos es el proceso de "migración", aunque a diferencia de él, existe un "mantenimiento" posterior. En realidad, la carga y mantenimiento de un almacén de datos es uno de los aspectos más delicados y que más esfuerzo requiere (alrededor de la mitad del esfuerzo necesario para implantar un almacén de datos), y, de hecho, suele existir un sistema especializado para realizar estas tareas, denominado Sistema de Extracción, Transformación y Carga (ETL por sus siglas en inglés). (6)

La construcción del ETL es responsabilidad del equipo de desarrollo del almacén de datos y se realiza específicamente para cada almacén de datos, aunque un ETL se puede construir realizando programas específicos.

1.3.5 Integración

El mecanismo de integración permite extraer la información de los sistemas fuentes (MES, ERP), luego transformarla y por último cargarla en el almacén de datos, donde corren las aplicaciones de inteligencia de negocio para el posterior análisis multidimensional de la información, lo cual permite el monitoreo y control mediante alertas e indicadores de rendimiento.

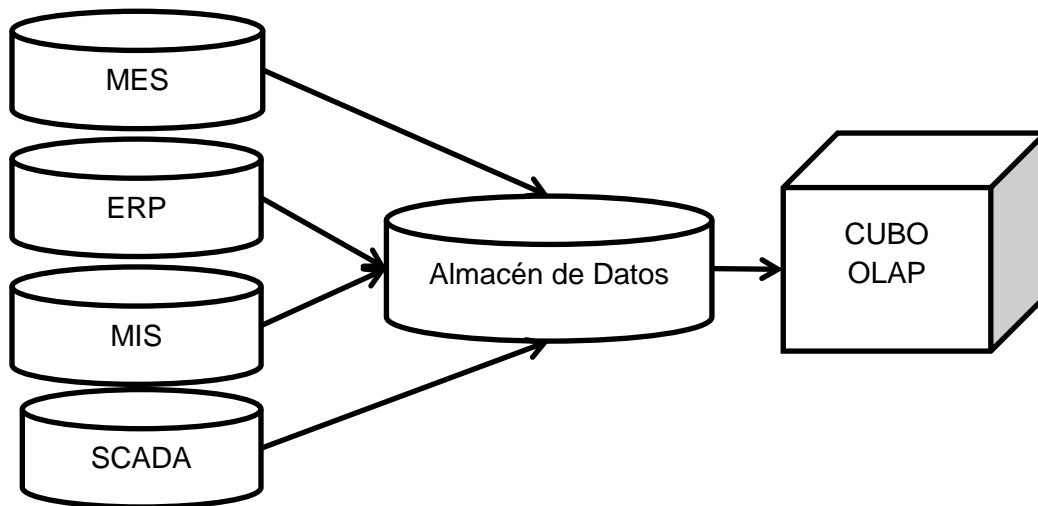


Figura 9. Integración de un sistema OLAP

En la presente investigación se tuvo en cuenta el análisis de herramientas que cumplieran con los requisitos para el desempeño de los estándares establecidos por la norma ISO 62264. A continuación se describen las herramientas utilizadas en la solución y la metodología aplicada.

1.4 Herramientas y Metodologías

Capítulo 1. Fundamentación Teórica.

1.4.1 SQL Server

SQL Server contiene "estudios" que ayudan en las tareas de programación y administración: SQL Server Management Studio y Business Intelligence Development Studio.

El Management Studio permite la administración del Motor de Base de Datos de SQL Server y la realización de soluciones de notificación, la administración de soluciones de Analysis Services implementadas, la ejecución de los paquetes de Integration Services y la administración de servidores de informes, así como los informes y modelos de informes de Reporting Services.

En Business Intelligence Development Studio, se desarrollan soluciones de Inteligencia de Negocio mediante proyectos de Analysis Services para desarrollar cubos, dimensiones y estructuras de minería; se crean proyectos de Reporting Services para la realización de informes y modelos de informes. Además se desarrollan proyectos de Integration Services para crear paquetes.

A continuación se describen las principales características de los productos de SQL Server utilizados en la propuesta de solución (8):

SQL Server Database Engine

El Database Engine (Motor de Base de Datos) es el servicio principal para almacenar, procesar y proteger datos. Este proporciona acceso controlado y procesamiento de transacciones rápido para cumplir con los requisitos de las aplicaciones consumidoras de datos más exigentes de la empresa.

SQL Server Analysis Services

Permite diseñar, crear y administrar estructuras multidimensionales que contienen datos detallados y agregados procedentes de varios orígenes de datos, como bases de datos relacionales, en un único modelo lógico unificado compatible con los cálculos integrados. También proporciona un análisis rápido, intuitivo y descendente de grandes cantidades de datos generados en este modelo de datos unificado, que se puede entregar a los usuarios en varios idiomas y monedas.

SQL Server Integration Services

Es una plataforma para la creación de soluciones empresariales de transformaciones e integración de datos. Integration Services sirve para resolver complejos problemas

Capítulo 1. Fundamentación Teórica.

empresariales mediante la copia o descarga de archivos, el envío de mensajes de correo electrónico como respuesta a eventos, la actualización de almacenes de datos, la limpieza y minería de datos, y la administración de objetos y datos de SQL Server. Los paquetes pueden funcionar por separado o conjuntamente con otros paquetes para hacer frente a las complejas necesidades de la empresa. Integration Services puede extraer y transformar datos de muchos orígenes distintos, como archivos de datos XML, archivos planos y orígenes de datos relacionales y, posteriormente, cargarlos en uno o varios destinos.

SQL Server Reporting Services

Es una plataforma de creación de informes que ofrece una completa funcionalidad de creación de informes para una gran variedad de orígenes de datos. Reporting Services contiene un completo conjunto de herramientas para crear, administrar y entregar informes, así como interfaces de programación de aplicaciones con las que los desarrolladores podrán integrar o extender el procesamiento de los datos y los informes en aplicaciones personalizadas.

1.4.2 Lenguaje de consulta MDX

MDX es un acrónimo de Multidimensional Query eXpression. Este lenguaje es en los sistemas OLAP el equivalente al Lenguaje de Consultas Estructurado (Structural Query Language, SQL por sus siglas en inglés) en los Sistemas Gestores de Base de Datos. MDX proporciona instrucciones del lenguaje de definición de datos (Data Definition Language, DDL por sus siglas en inglés); instrucciones del lenguaje de manipulación de datos (Data Manipulation Language, DML por sus siglas en inglés); instrucciones del lenguaje de secuencias de comandos para administrar el ámbito, el contexto y el control de flujo en secuencias de comandos MDX; operadores y funciones para la manipulación de datos recuperados a partir de objetos multidimensionales; y la posibilidad de ampliar el lenguaje MDX con funciones definidas por el usuario. (6)

Todo lo anteriormente expuesto significa que es el lenguaje a través del cual se puede explotar la información que reside en los motores OLAP y satisfacer las consultas analíticas ya que tiene funciones y fórmulas que lo hacen muy potente.

1.4.3 Metodología

Existen muchas metodologías de diseño y construcción de un almacén de datos (DW). Cada fabricante de software de inteligencia de negocios busca imponer una metodología con sus productos. Sin embargo, se imponen entre la mayoría dos metodologías, la de Kimball y la de

Capítulo 1. Fundamentación Teórica.

Inmon. Para comprender la mayor diferencia entre estas dos metodologías, debemos explicar el concepto de un Datamart. Un Datamart es un repositorio de información, similar a un DW, pero orientado a un área o departamento específico de la organización (por ejemplo Calidad, Inventario, RRHH, etc.), a diferencia del DW que cubre toda la organización, es decir la diferencia fundamental es su alcance.

Por otra parte, la metodología de Inmon se basa en conceptos bien conocidos del diseño de bases de datos relacionales; la metodología para la construcción de un sistema de este tipo es la habitual para construir un sistema de información, utilizando las herramientas habituales, al contrario de la de Kimball, que se basa en un modelado dimensional (no normalizado) (9).

Pensamos que la metodología más acorde a los negocios de nuestra región es la de Kimball, por cuanto proporciona un enfoque de menor a mayor, muy versátil, y una serie de herramientas prácticas que ayudan a la implementación de un DW. Es acorde a nuestras empresas porque se pueden implementar pequeños Datamarts en áreas específicas de las mismas, con pocos recursos y de poco irlos integrándolos en un gran almacén de datos.

La metodología de Kimball proporciona una base empírica y metodológica adecuada para las implementaciones de almacenes de datos pequeños y medianos, dada su gran versatilidad y su enfoque ascendente, que permite construir los almacenes en forma escalonada. Además presenta una serie de herramientas, tales como planillas, gráficos y documentos, que proporcionan una gran ayuda para iniciarse en el ámbito de la construcción de un almacén de datos.

La metodología se basa en lo que Kimball denomina Ciclo de Vida Dimensional del Negocio (Business Dimensional Lifecycle por su término en inglés). Este ciclo de vida del proyecto de almacén de datos, está basado en cuatro principios básicos:

- 1) Centrarse en el negocio: Hay que concentrarse en la identificación de los requerimientos del negocio y su valor asociado, y usar estos esfuerzos para desarrollar relaciones sólidas con el negocio, agudizando el análisis del mismo y la competencia consultiva de los implementadores.
- 2) Construir una infraestructura de información adecuada: Diseñar una base de información única, integrada, fácil de usar, de alto rendimiento donde se reflejará la amplia gama de requerimientos de negocio identificados en la empresa.
- 3) Realizar entregas en incrementos significativos: crear el almacén de datos (DW) en incrementos entregables en plazos de 6 a 12 meses. Hay que usar el valor de negocio

Capítulo 1. Fundamentación Teórica.

de cada elemento identificado para determinar el orden de aplicación de los incrementos. En esto la metodología se parece a las metodologías ágiles de construcción de software.

- 4) Ofrecer la solución completa: proporcionar todos los elementos necesarios para entregar valor a los usuarios de negocios. Para comenzar, esto significa tener un almacén de datos sólido, bien diseñado, con calidad probada, y accesible. También se deberá entregar herramientas de consulta, aplicaciones para informes y análisis avanzado, capacitación, soporte, sitio web y documentación.

En la siguiente figura se ilustra el ciclo de vida dimensional del negocio donde se distinguen las principales actividades de la metodología que se realizan durante el desarrollo de la propuesta de solución. Se distinguen las actividades centrales como las más importantes de la metodología y en las cuales se centra el diseño de la solución OLAP.

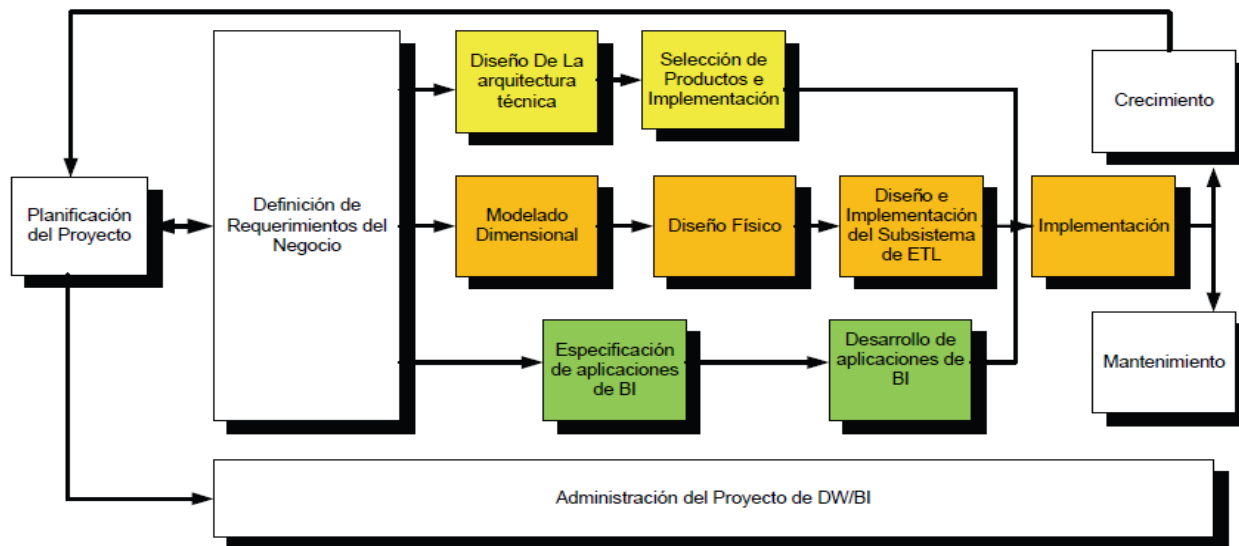


Figura 10. Tareas de la Metodología de Kimball

1.5 Conclusiones

En este capítulo se realizó un estudio de los fundamentos teóricos que soportan la investigación científica, relacionados con el objeto de estudio del presente trabajo. Se abordaron, a grandes rasgos, temas como la generación de energía eléctrica, el nivel de automatización existente en plantas de generación de energía distribuida y la Inteligencia de Negocio para soportar el monitoreo y control de los sistemas de gestión empresarial basados en la Norma ISO 62264. Además, se expusieron los elementos de la tecnología OLAP, teniéndose como resultado la selección del sistema MOLAP. Se decide utilizar SQL Server

Capítulo 1. Fundamentación Teórica.

2008 R2 por ser compatible con el nivel de automatización dado en la planta GEDICH y los productos de la plataforma de SQL Server así como la metodología de Kimball para guiar el proceso de desarrollo.

2 Capítulo 2. Propuesta de solución

En este capítulo se abordan todos los elementos de la propuesta desde la concepción de la solución, la descripción de los requisitos funcionales y no funcionales, el modelado dimensional del Cubo OLAP, el diseño físico del Almacén de Datos y la implementación de la solución.

2.1 Descripción de la Infraestructura en la capa superior de Control

En GEDICH existe una infraestructura que sirve de base para la aplicación de Inteligencia de Negocio, que actualmente cuenta con los siguientes servidores:

Servidores	Características y aplicación
Controladores de Dominio	Permiten sostener toda la Infraestructura de servicios de la red, la seguridad, el control de usuarios de la capa de aplicaciones, la mensajería y la navegación, basados en Windows 2003 Server Ent. Ed.
Servidores de Proceso SIMATIC WinCC 7.0	Calcula algunas variables de integración que devuelve a la capa superior. Sostiene los servicios de WebNavigator y Connectivity Pack para monitorear el proceso sobre web y el acceso a la capa de datos del sistema SCADA respectivamente, en una arquitectura cliente servidor, con el servidor en la red tecnológica obteniendo las variables de proceso y mediante la definición de PowerPacks. El Servidor Cliente de WinCC es el encargado de llegar a través de un Firewall de proceso al Servidor de proceso y suministrar una consola Web para los usuarios de la red de gestión.
Firewall de Proceso	Un servidor MS ISA Server encargado de proteger la red y accesos no autorizados.
Servidor de Bases de Datos	Servidor MS SQL Server 2008 R2 Empresarial que constituye el motor de bases de datos de todas las aplicaciones OLTP, que servirán en la operación y el control de la planta, y además

Capítulo 2. Propuesta de solución

	constituyen las fuentes de datos del Cubo MOLAP de GEDIMES.
Servidor con el Servicio de Análisis	Servidor MS SQL Server 2008 R2 que soporta el Cubo MOLAP de la aplicación GEDIMES.
Servidor de Aplicaciones	Soporta la capa web de GEDIMES como interfaz para las operaciones de sus sistemas básicos y la consola de inteligencia, donde se muestran de forma gráfica los KPI y las curvas de análisis.
Servidores de Reportes e Integración	Implementan los reportes diseñados para GEDIMES y las tareas del flujo de integración que ajustan las actividades de integración en la plataforma, basado en MS Reporting Server and Integration Services, mediante el estándar B2MML, recomendado en la ISO/IEC 62264.

Tabla 1. Relación de Servidores en la Plataforma GEDIMES

Es importante relacionar las aplicaciones que deben formar parte del entorno operacional de las plantas para poder obtener el Cubo MOLAP.

Aplicaciones	Características fundamentales
Sisgesmant	Aplicación MS Windows/ MS ASP.Net Cliente servidor para la gestión de los mantenimientos a partir del patrimonio de una planta.
SigereRH	Solución MS Windows para la gestión de los recursos humanos en las plantas, con un importante módulo de presencia en automático que permite obtener el control de personal por cada turno de trabajo.
GEDIMES Laboratorio	Aplicación MS ASP.Net MVC que permite la gestión de las pruebas de calidad de las materias de Servicios, según arquitectura de la IIEC/ISO 62264.
GEDIMES Inventario	Aplicación MS ASP.Net MVC que permite la gestión de los

Capítulo 2. Propuesta de solución

	Inventarios de los Insumos al proceso de producción en la planta, según arquitectura de la IIEC/ISO 62264.
GEDIMES Producción	Aplicación MS ASP.Net MVC que permite la gestión el control de las operaciones de producción en la planta, según arquitectura de la IIEC/ISO 62264.
GEDIMES BI	Aplicación MS ASP.Net MVC que contiene la consola de Supervisión del negocio a nivel de planta para los que toman decisiones, demás cuenta con el monitor de KPI, la interfaz de la Tabla de Pívor de GEDIMES y el seguidor gráfico del proceso de adaptación de la planta.
Otras	Conjunto de aplicaciones que dan respuesta a otros subsistemas de la planta, paralelo al productivo.

Tabla 2. Relación de OLTP que alimenta el Cubo MOLAP GEDIMES y su Aplicación Consola GEDIMES BI

La tabla anterior permite observar que la solución que se desarrolló forma parte de un sistema para el control de la producción de energía distribuida. A continuación se exponen los artefactos más importantes que describen la lógica de la solución, con el objetivo de mostrar una idea clara de las fuentes de datos de donde se nutrirá la solución MOLAP, los mismos condicionan uno de los extremos en los flujos de integración, que responden a la dinámica de la plataforma de gestión, basados en los múltiples procesos que intervienen dentro de GEDIMES.

2.2 GEDIMES

GEDIMES es una solución MES, basada en una plataforma Web, que gestiona el proceso de producción, el control de los inventarios, las pruebas de calidad. Además, ofrece una interfaz para las solicitudes de órdenes de trabajo y sus respuestas, que interactúa con un sistema de gestión de mantenimientos, perteneciente a una solución de terceros que se integra a la plataforma. Su arquitectura está diseñada bajo la norma ISO/IEC 62264, que permite un alto nivel de estandarización e integración entre los subsistemas de control de la producción al cual pertenece la plataforma y los sistemas empresariales referentes al módulo de Inteligencia de Negocio.

Capítulo 2. Propuesta de solución

A continuación se muestra el diagrama de clases de GEDIMES, que constituye un artefacto de entrada para el modelo del Cubo OLAP.

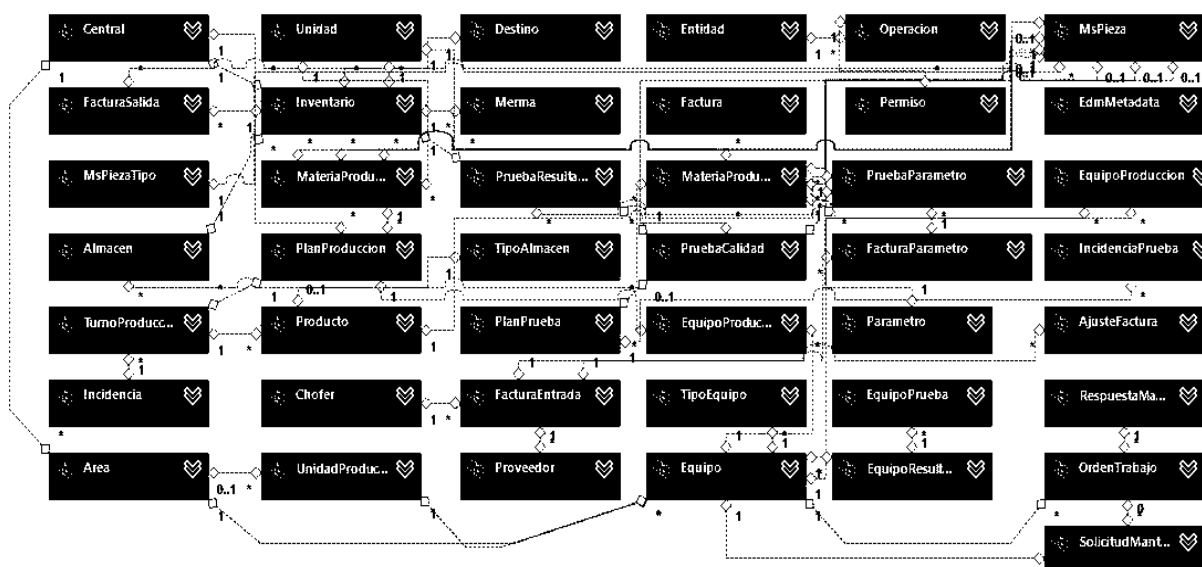


Figura 11. Diagrama de Clases del Sistema GEDIMES.

Este diagrama está dividido en cuatro módulos básicos (Producción, Mantenimiento, Inventario, y Laboratorio) que responden a los procesos críticos de una planta. GEDIMES permite el análisis de forma estática de la información, así como tomar decisiones a corto plazo basadas en la planificación de la producción y la gestión de los recursos. El modelo no permite el análisis de la información en tiempo real, ya que su base de datos es OLTP, por lo que solo procesa solicitudes transaccionales que no responden al control de la producción de una manera dinámica, compleja y en cortos instantes de tiempo.

Para dar solución a este problema se implementó un Cubo MOLAP basado en la información almacenada en un conjunto de bases de datos OLTP donde GEDIMES es el eje central.

2.3 Requisitos

2.3.1 Requisitos del negocio

La definición de los requisitos del negocio (RN) es en gran medida un proceso de entrevistar al personal de negocio y técnico, pero siempre conviene tener un poco de preparación previa. Se debe aprender tanto como se pueda sobre el negocio, los competidores, la industria y los clientes del mismo. Hay que leer todos los informes posibles de la organización; rastrear los documentos de estrategia interna; entrevistar a los empleados, analizar lo que se dice en la

Capítulo 2. Propuesta de solución

prensa acerca de la organización, la competencia y la industria. Se deben conocer los términos y la terminología del negocio (7).

A continuación se describen los requisitos del negocio del Almacén de Datos.

RN1. Consultar información mediante tablas Pívor desde un cliente OLAP para el análisis multidimensional de la información.

RN2. Proporcionar información sobre el estado real de la planta basado en los indicadores claves de rendimiento.

RN3. Notificar mediante el análisis lógico algebraico los problemas detectados por el sistema.

RN4. Generar informe acerca del estado de la planta.

RN5. Generar informe acerca del funcionamiento modular de la planta.

RN6. Gestionar los datos del almacén de datos.

De acuerdo a la metodología de Kimball se deben relacionar los temas analíticos y los requisitos especificados. Los temas analíticos son las operaciones de producción de GEDICH.

En la siguiente tabla se muestra dicha relación:

No.	Tema Analítico	Requisitos	Prioridad
1	Producción	RN1, RN2, RN3, RN4, RN5, RN6	Alta
2	Mantenimiento	RN1, RN2, RN3, RN4, RN5, RN6	Alta
3	Laboratorio	RN1, RN2, RN3, RN4, RN5, RN6	Alta
4	Inventario	RN1, RN2, RN3, RN4, RN5, RN6	Alta

Tabla 3. Operaciones basadas en la implementación de los requisitos.

2.3.2 Restricciones del negocio

Las restricciones de negocio (RTN) no son más que condiciones que deben satisfacerse o políticas que deben cumplirse, por lo que éstas regulan el negocio. Luego de realizar un análisis minucioso del negocio se identificaron las reglas de negocio, a continuación se muestran algunas de las reglas identificadas.

Usabilidad

Capítulo 2. Propuesta de solución

RTN1. Proporcionar las características mínimas de hardware que permitan el correcto funcionamiento de la solución.

Las especificaciones de hardware mínimas de cada nodo se muestran a continuación, El funcionamiento ideal del sistema está conformado por la arquitectura representada en la siguiente figura.

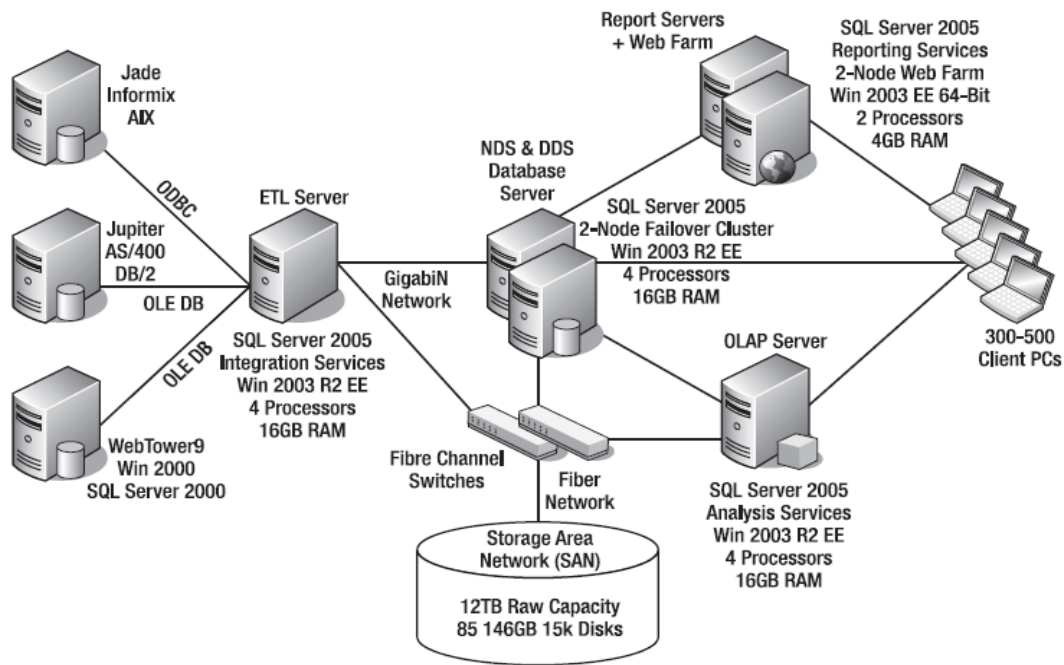


Figura 12. Arquitectura física de hardware para implementar el sistema.

RTN2. Instalar en los servidores las aplicaciones necesarias para el correcto funcionamiento del sistema.

En los servidores deben estar instaladas las siguientes herramientas:

- Microsoft Windows Server 2008 R2 Standard Edition o una versión posterior.
- Microsoft SQL Server 2008 R2 Enterprise Edition o una versión posterior.
- SQL Server 2008 Report Builder 2.0.
- Visual Studio 2008 BI
- Office SharePoint Server 2010.
- Office 2010 Professional Edition o una versión posterior, incluyendo Excel 2010 y Visio 2010.
- SQL Server 2008 Data Mining Add-ins para Office 2010.

Capítulo 2. Propuesta de solución

Confiabilidad

- RTN3. Proteger la información de publicidad y de acceso no autorizado a la misma.
- RTN4. Reportar problemas de conectividad hacia las fuentes de datos.

Seguridad

- RTN5. No hacer pública información privada del proyecto Gedimes BI.
- RTN6. No divulgar los datos de Gedimes y a los mismos sólo puede acceder personal autorizado.

SopORTE

- RTN7. Garantizar la instalación y ejecución de la aplicación de inteligencia de negocio en un entorno controlado y con restricciones de acceso.

Restricciones de diseño

- RTN8. Desarrollar a través de la metodología de Kimball cada una de las fases basadas en el ciclo dimensional del negocio.
- RTN9. Desarrollar utilizando SQL Server Business Intelligence Development Studio como entorno integrado de desarrollo y MDX como lenguaje de consulta de datos.

Estándares aplicables

- RTN10. Diseñar la integración del sistema de planificación y el sistema de control de la producción a partir de las especificaciones de la norma ISO 62264.

2.4 Modelado Dimensional (Cubo MOLAP)

El Cubo MOLAP de GEDIMES es una base de datos multidimensional, donde el almacenamiento físico de los datos se realiza en un vector multidimensional que permite el análisis y la obtención de recetas e indicadores claves de rendimiento que marcan el estado actual de la planta.

La estructura del diseño del cubo está formada por hechos y dimensiones. Los **hechos** son tablas que almacenan las medidas de las actividades críticas del sistema y las tablas de parámetros en función de los cuales se analizan los datos de manera jerárquica se conocen como **dimensiones**. Las formas que tienen estos conjuntos de hechos y dimensiones son llamadas almacenes de datos.

Capítulo 2. Propuesta de solución

Los hechos críticos identificados son Producción, Mantenimiento, Laboratorio e Inventario, que contienen atributos que miden el proceso. Estos son llamados métricas o medidas, que al aplicar entre ellos operadores algebraicos se convierten en Indicadores Claves de Rendimiento (KPI), que miden el nivel del desempeño del proceso utilizado para cuantificar objetivos que reflejan el rendimiento de la planta. Los hechos en GEDIMES tienen relaciones directas con las dimensiones. En la siguiente tabla se especifica dicha relación.

Proceso de Negocio	Dimensiones						
	Tiempo	Equipo	Materia y Pieza	RRHH	Prueba	Factura	Orden Trabajo
Producción	X		X	X			
Mantenimiento	X	X	X	X			X
Laboratorio			X	X	X		
Inventario			X			X	

Tabla 4. Matriz de procesos/dimensiones (Bus Matrix)

La matriz de procesos - dimensiones tiene en sus filas y columnas los procesos de negocio y las dimensiones identificadas respectivamente. Cada **X** en la intersección de las filas y columnas significa en qué **proceso de negocio** en la fila seleccionada se identifican las **dimensiones** propuestas.

El diseño es basado en una estrella simple porque no hay caminos alternativos en las dimensiones. En general los almacenes de datos contienen información redundante, ya que la desnormalización permite un mayor rendimiento en el análisis. En la siguiente figura se muestra el resultado del diseño físico del Cubo MOLAP de GEDIMES:

Capítulo 2. Propuesta de solución

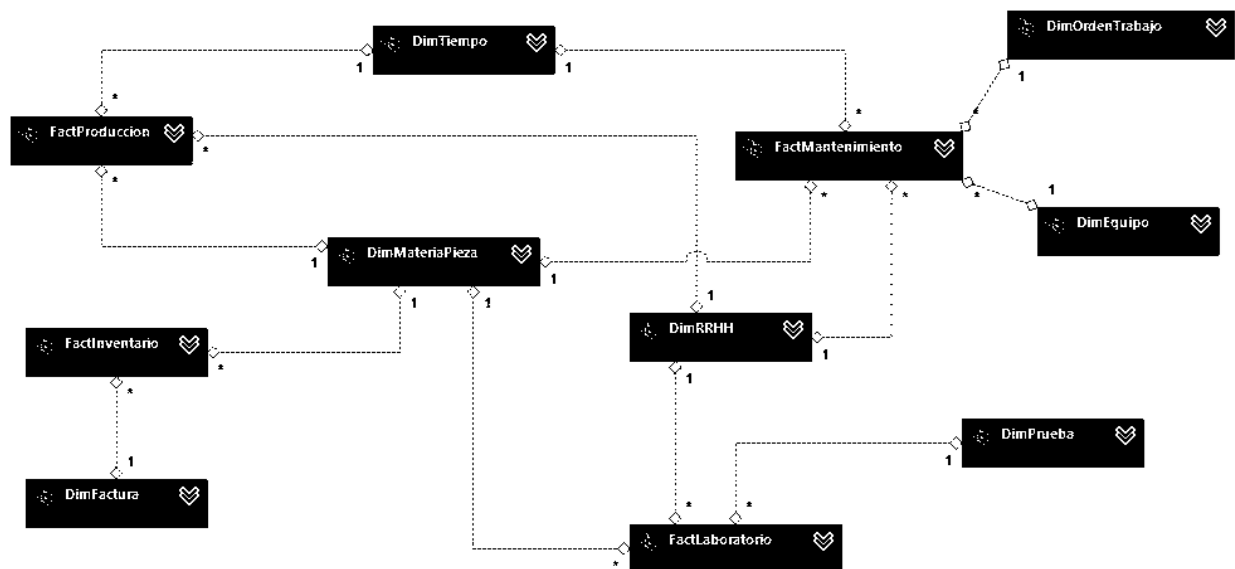


Figura 13. Diseño Multidimensional del Cubo MOLAP de GEDIMES.

2.5 Diseño Físico del Almacén de Datos

A continuación se muestra en las tablas (Tabla 5, Tabla 6, Tabla 7, Tabla 8) una descripción detallada acerca de las dimensiones que se asocian a los hechos críticos del sistema. En cada tabla hecho la primera columna hace referencia a las llaves de las dimensiones relacionadas y a partir de las cuales se comprenden las medidas que intervienen en el proceso.

2.5.1 Diseño físico de los hechos

2.5.1.1 Diseño físico de la tabla Producción

Columna	Tipo Dato	Descripción	Ejemplo
ProduccionId	int	Llave del hecho	5
RrhhlId	int	Llave de la dimensión RRHH	124
MateriaPiezaId	int	Llave de la dimensión MateriaPieza	45
Tiempold	int	Llave de la dimensión Tiempo	20120311
CantAguaProducida	int	Cantidad de agua producida	4655
CantArranque	int	Cantidad de arranque motor	125

Capítulo 2. Propuesta de solución

ConsumoMateria	double	Consumo de la materia de servicio	4656
EnergiaGenerada	double	Cantidad de energía generada	654
ProduccionLodos	double	Cantidad de producción de lodos	7895
TotalHorasTrabajadas	time	Total de horas trabajadas	12

Tabla 5. Hecho Producción.

En la siguiente figura se muestra cómo se produce la relación entre el hecho producción y las dimensiones RRHH, Tiempo y MateriaPieza. Se puede interpretar que el producto se generó con un recurso humano, una materia de servicio y un tiempo determinado.

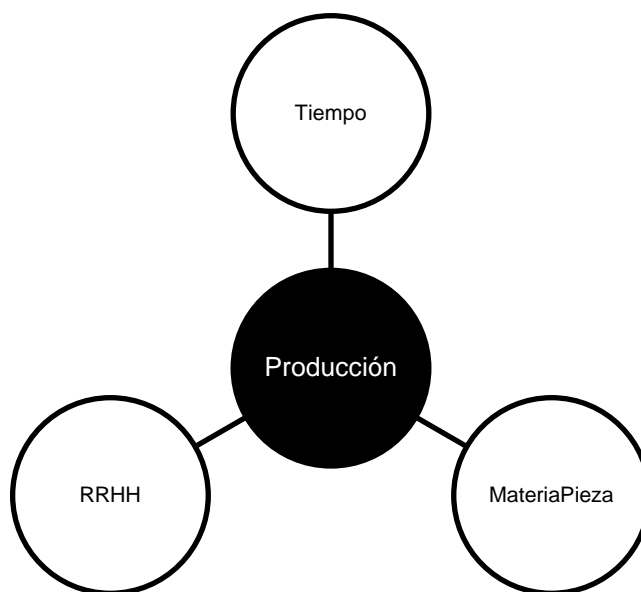


Figura 14. Relación multidimensional del hecho Producción

2.5.1.2 Diseño físico de la tabla Mantenimiento

Columna	Tipo Dato	Descripción	Ejemplo
ProduccionId	int	Llave de la dimensión	5
Equipold	int	Llave de la dimensión Patrimonio	124
Rrhhd	int	Llave de la dimensión RRHH	45
MateriaPiezald	int	Llave de la dimensión Materias y Piezas	20120311

Capítulo 2. Propuesta de solución

OrdenTrabajold	int	Llave de la dimensión Orden de Trabajo	4655
Tiempold	int	Llave de la dimensión Tiempo	125
CantEquiposIndisponibles	int	Cantidad de equipos indisponibles	4656
CantMantenimientoPlanificado	int	Cantidad de mantenimiento planificado	654
CantMantenimientoRealizado	int	Cantidad mantenimiento realizado	7895
CantOrdenTrabPendientes	int	Cantidad órdenes trabajo pendientes	12
CantOrdenTrabTotal	int	Cantidad órdenes trabajo total	45
CantFallas	int	Cantidad de fallas	65
NoIntervencionPreventiva	double	Número de Intervención preventiva	25
NoItem	double	Número de Ítem	12
TiempoEjecucion	time	Tiempo de ejecución	12:24:12
TiempoIntervencionCorrectiva	time	Tiempo intervención correctiva	13:24:12
TiempoIntervencionPreventiva	time	Tiempo intervención preventiva	09:24:12
TiempoOperacion	time	Tiempo de operación	01:24:12
TiempoParadoPiezas	time	Tiempo parado por piezas	02:24:12
TiempoPlanificado	time	Tiempo planificado	10:24:12
TiempoPromDiagnostico	time	Tiempo promedio para un diagnóstico	08:24:12
TiempoPromSuministro	time	Tiempo promedio para Suministro	16:24:12
TiempoPromTrabajo	time	Tiempo promedio de Trabajo	22:24:12
TotalEquipos	int	Total de equipos	6455
TotalHorasHombre	time	Total horas hombre	21:24:12

Capítulo 2. Propuesta de solución

TotalHorasTrabajadas	time	Total horas trabajadas	15:24:12
----------------------	------	------------------------	----------

Tabla 6. Hecho Mantenimiento.

En este caso, el hecho significa que los mantenimientos son realizados por un **recurso humano** a un **equipo**, usando una **materia de servicio**, solicitado por una **orden de trabajo** en un **tiempo** determinado.

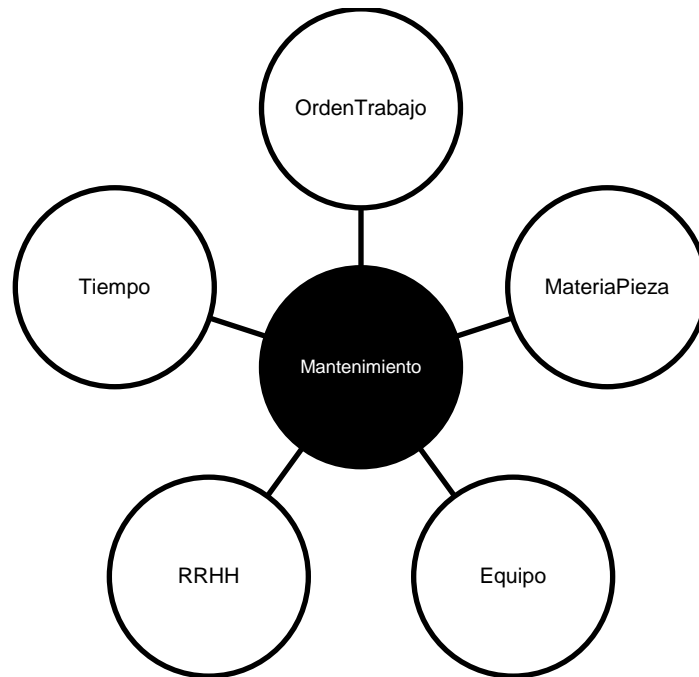


Figura 15. Relación multidimensional del hecho Mantenimiento

2.5.1.3 Diseño físico de la tabla Laboratorio.

Columna	Tipo Dato	Descripción	Ejemplo
Laboratoriold	int	Llave del hecho	12
Rrhhd	int	Llave de la dimensión RRHH	32
MateriaPiezald	int	Llave de la dimensión MateriaPieza	52
Pruebald	int	Llave de la dimensión Prueba	452
%Ejecucion	double	Porciento de Ejecución	96
%RealizaciónTecnica	double	Porciento de Realización Técnica	84

Capítulo 2. Propuesta de solución

%Satisfaccion	double	Porciento de Satisfacción	97
CantidadAditivoQuímico	int	Cantidad de Aditivo Químico	656
CantidadPruebaSolicitada	int	Cantidad de Pruebas Solicitadas	5335
CantidadPruebaRealizada	int	Cantidad de Pruebas Realizadas	652
NoConformidad	int	Cantidad de No Conformidades	3
TiempoEjecucion	time	Tiempo de Ejecución	10:23:52

Tabla 7. Hecho Laboratorio.

En este caso, el hecho Laboratorio resuelve una **prueba** que fue realizada por un **recurso humano** a una **materia de servicio**. Estas pruebas no están dimensionadas en el tiempo debido a que dependen de un análisis químico a cierta cantidad de materia.

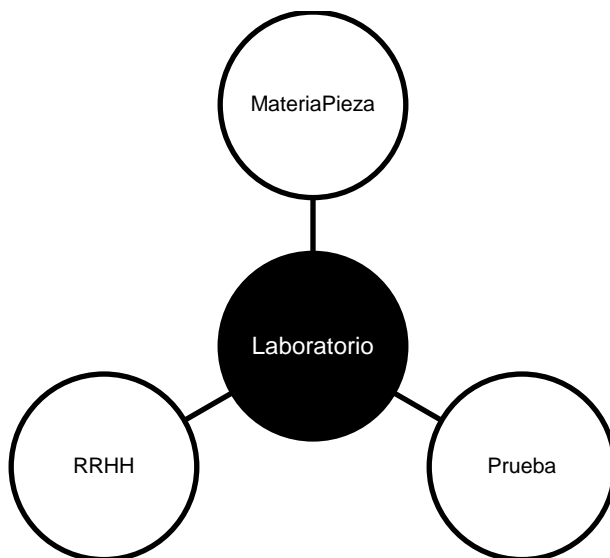


Figura 16. Relación multidimensional del hecho Laboratorio

2.5.1.4 Diseño físico de la tabla Inventario

Columnas	Tipo de Dato	Descripción	Ejemplo
Inventariold	int	Llave del hecho	25
MateriaPiezald	int	Llave de la dimensión MateriaPieza	53

Capítulo 2. Propuesta de solución

FacturaId	int	Llave de la dimensión Factura	45
CantidadPerdida	int	Cantidad de Pérdidas	1352
CantidadStock	int	Cantidad en Stock	5421
CapacidadTotal	int	Capacidad total del almacén	6548
FlujoExpedicion	time	Tiempo en expedir una materia o pieza	01:12:25
FlujoRecepcion	time	Tiempo en recibir una materia o pieza	00:25:45
PrecioTotal	decimal	Precio total de la materia o pieza	3265.00

Tabla 8. Hecho Inventario.

Para este evento, refleja la dinámica del proceso inventario, donde cierta cantidad de recurso entra o salen del proceso por una factura de proveedor o solicitud a la producción.



Figura 17. Relación multidimensional del hecho Inventario

A continuación se muestra una descripción de las dimensiones identificadas en nuestro modelo, estas siguen una estructura jerárquica que permite obtener las medidas a un mayor nivel de detalle.

Capítulo 2. Propuesta de solución

2.5.2 Diseño físico de las dimensiones

2.5.2.1 Diseño físico de la tabla Equipo

Columna	Tipo de Dato	Descripción	Ejemplo
Patrimoniold	int	Llave de la dimensión.	25
PatrimonioldAlternativo	string	Llave de la tabla fuente.	EC-05
Area	string	Área donde se ubica el equipo	Batería 01
Estado	string	Estado del Equipo.	Excelente
Central	string	Área donde se ubica el patrimonio	GEDICH
Tipo	string	Tipo de Equipo	Generador
Identificador	string	Nombre del Equipo	G-21

Tabla 9. Dimensión Equipo.

Jerarquía de la Dimensión Equipo:

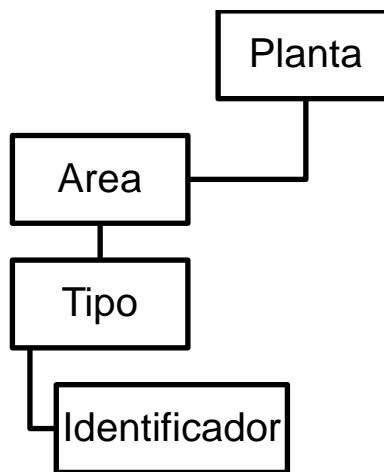


Figura 18. Jerarquía de la Dimensión Equipo

2.5.2.2 Diseño físico de la tabla RRHH

Columna	Tipo de Datos	Descripción	Ejemplo
Rrhhld	int	Llave de la dimensión	25

Capítulo 2. Propuesta de solución

RrhIdAlternativo	string	Llave de la tabla fuente	T-56
Nombre	string	Nombre del Trabajador	Roberto
Primer Apellido	string	Apellidos del Trabajador	Gómez
Segundo Apellido	string	Apellidos del Trabajador	Pérez
Nombre Completo	string	Nombres y Apellidos	Roberto Pérez Gómez
Sexo	string	Sexo del Trabajador	Masculino
Dirección	string	Dirección del Trabajador	San Rafael #34 e/ ...
Municipio	string	Municipio de la provincia	Centro Habana
Provincia	string	Provincia del país	La Habana
Cargo	string	Cargo del Trabajador	Administrador

Tabla 10. Dimensión RRHH.

Jerarquía de la Dimensión RRHH:

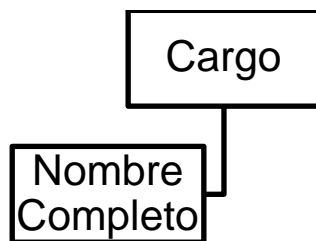


Figura 19. Jerarquía de la Dimensión RRHH

2.5.2.3 Diseño físico de la tabla MateriaPieza

Columna	Tipo Dato	Descripción	Ejemplo
MateriaPiezald	int	Llave de la dimensión	1, 2, 3...
MateriaPiezaldAlternativo	string	Llave de la tabla fuente	L8556

Capítulo 2. Propuesta de solución

Descripción	string	Descripción de la materias o piezas	Aceite
Estado	string	Estado actual de los materiales o piezas	Disponible
Codigo	string	Proveedor de los materiales y piezas	M-52
Tipo	string	Define si es materia o pieza	Materia
Unidad	string	Unidad de la materia o pieza	Libra

Tabla 11. Dimensión MateriaPieza.

Jerarquía de la Dimensión Materia y Pieza

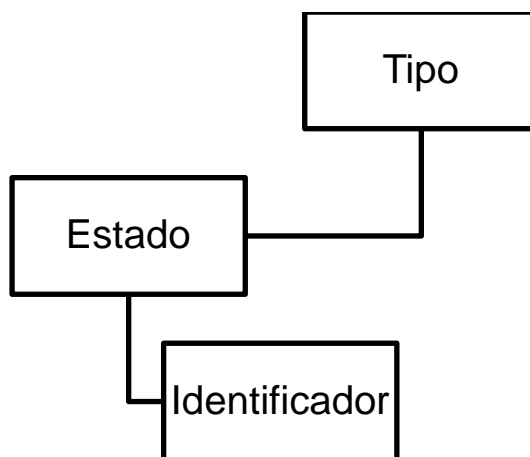


Figura 20. Jerarquía de la Dimensión MateriaPieza

2.5.2.4 Diseño físico de la tabla Prueba

Columna	Tipo Dato	Descripción	Ejemplo
Pruebald	int	Llave de la dimensión	16
PruebaldAlternativo	string	Llave de la tabla fuente	P465
TipoPrueba	string	Tipos de pruebas para laboratorio	Agua cruda
Descripcion	string	Descripción de la prueba de calidad	Nivel de agua cruda

Tabla 12. Dimensión Prueba.

Jerarquía de la Dimensión Prueba:

Capítulo 2. Propuesta de solución

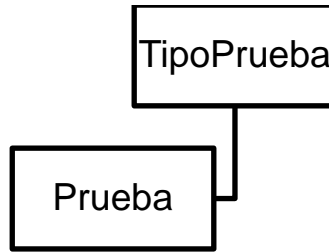


Figura 21. Jerarquía de la Dimensión Prueba

2.5.2.5 Diseño físico de la tabla OrdenTrabajo

Columna	Tipo de Datos	Descripción	Ejemplo
OrdenTrabajold	int	Llave de la dimensión	25
OrdenTrabajoldAlternativo	string	Llave de la tabla fuente	OT-56
Codigo	string	Código de la orden	652
Descripcion	string	Descripción de la avería	Avería en tanque C001
Taller	string	Descripción de la ubicación	Taller01
TipoMantenimiento	string	Clasificación del mantenimiento	Planificado
TiempoEjecucion	time	Tiempo que se demora la ejecución	02:45:52
Estado	string	Estado de la avería a solucionar	Grave

Tabla 13. Dimensión OrdenTrabajo

Jerarquía de la Dimensión Orden Trabajo:

Capítulo 2. Propuesta de solución

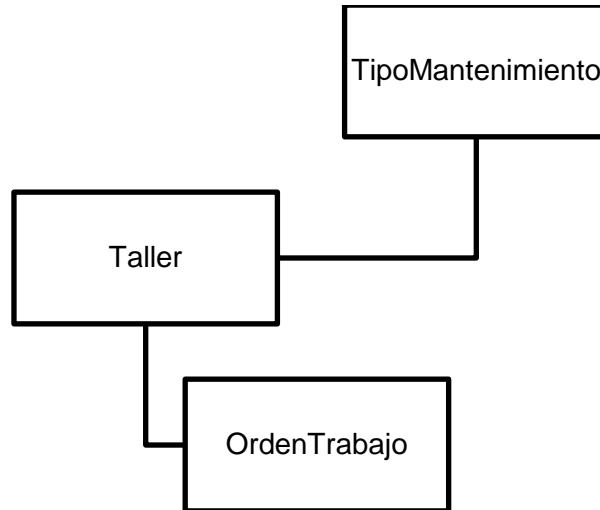


Figura 22. Jerarquía de la Dimensión OrdenTrabajo

2.5.2.6 Diseño físico de la tabla Factura

Columna	Tipo de Datos	Descripción	Ejemplo
FacturaId	int	Llave de la dimensión	25
FacturaIdAlternativo	string	Llave de la tabla fuente	F568
Cantidad	string	Cantidad en inventario	652
Codigo	string	Código de la factura	F456
Lote	string	Lote de la factura	L568
PrecioUnitario	string	Precio por unidad	325.00
TipoFactura	string	Tipo de factura para inventario	Entrada

Tabla 14. Dimensión Factura.

Jerarquía de la Dimensión Factura:

Capítulo 2. Propuesta de solución

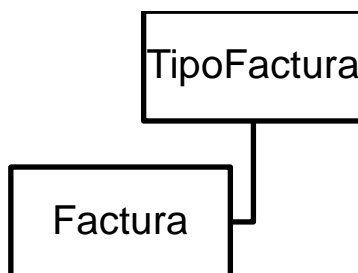


Figura 23. Jerarquía de la Dimensión Factura

2.6 Indicadores Claves de Rendimiento

Una vez mostrado el diseño físico del Cubo MOLAP y sus métricas se definieron los Indicadores Claves de Rendimiento (KPI), que definen en el espacio multidimensional, el estado del comportamiento del proceso en tiempo real. Esta actividad se realiza en paralelo con el diseño del mecanismo ETL que se explica más adelante. Es importante reflejar que los KPI son funciones aritméticas cuyos argumentos están constituidos por las métricas obtenidas del proceso. Por tanto estos indicadores son una imagen del estado de la gestión del proceso de producción de energía eléctrica en la planta o la empresa.

Un KPI posee cuatro atributos fundamentales: valor, objetivo, estado y tendencia. El valor está asociado al resultado actual que toma el KPI, describiendo la situación más actualizada del estado de la planta. El valor objetivo o simplemente objetivo, no es más que el valor definido por la organización para cuando el proceso alcanza una realización óptima de las métricas que definen el KPI mediante una relación determinada. El estado y la tendencia son variables cuyos valores a alcanzar pertenecen a la triada (-1, 0, 1) significando una diferencia lógica entre los valores de estado y el objetivo respectivamente.

En el Manual de Gestión para la Generación Distribuida de Electricidad en Cuba (MAGEST) (1) se especifican los KPI para cada una de las actividades críticas del proceso de generación. A continuación se describen los indicadores por cada una de las actividades en la siguiente tabla.

Nombre	Fórmula	Descripción
Consumo Específico de Combustible	$CE = \frac{\text{Consumo Comb.}}{\text{Energía Generada}}$	Es la cantidad en gramos de combustible que se consume por cada kilowatts generado en un intervalo de

Capítulo 2. Propuesta de solución

		tiempo.
Consumo Específico de Lubricante Motor	$CE_{ac} = \text{Consumo Aceite Motor} / \text{Energía Generada}$	Es el lubricante que se consume en gramos por cada kilowatts generado en un intervalo de tiempo.
Consumo Específico de Refrigerante	$CE_{ref} = \text{Consumo Refrigerante} / \text{Energía Generada}$	Es el refrigerante que se consume en litros por cada kilowatts generado en un intervalo de tiempo.
Consumo Específico de Agua	$CE_{agua} = \text{Consumo Agua} / \text{Energía Generada}$	Es el agua que se consume en litros por cada kilowatts generado en un intervalo de tiempo.
Consumo Específico de Lubricantes Auxiliares	$CE_{acx(x)} = \text{Consumo Lubricante (x)} / \text{Horas de Trabajo}$	Es el lubricante que se consume en litros por cada hora de trabajo del equipo auxiliar.
Consumo Específico de Productos Químicos	$CEPQ(x) = \text{Consumo (x)} / \text{Litros de Líquidos}$	Es la cantidad en litros de producto químico consumido por litro del fluido a tratar.
Porcentaje de órdenes de trabajo pendientes	$\%OTP(x) = \text{Órdenes de Trabajo Pendiente} / \text{Órdenes de Trabajo Total}$	Es la relación entre las órdenes de trabajo pendiente y las órdenes de trabajo totales llevados a por cientos
Disponibilidad	$D = (\text{Potinstalada} - \text{Potindistotal}) / (\text{Potinstalada}) \times 100$	Es la relación entre la potencia total instalada y la potencia real que existe en

Capítulo 2. Propuesta de solución

		<p>el período dado. La potencia que no puede cumplir su función por mantenimiento, avería o causas externas forma parte de la diferencia entre la instalada y real en el período analizado.</p>
<p>Tiempo promedio entre fallos</p>	$TPEF = NOIT * HROP / \sum NTMC$	<p>Relación entre el producto del número de ítems por sus tiempos de operación y el número total de fallas detectadas, en esos ítems en el período observado.</p>
<p>Tiempo promedio para mantenimiento</p>	$TPPM = \sum HTMP / NTMC$	<p>Relación entre el tiempo total de la intervención preventiva y el número total de la intervención preventiva en esos ítems, en el período observado.</p>
<p>Tiempo promedio para la reparación</p>	$TPPR = TPPD + TPPS + TPPT$	<p>Relación entre la sumatoria de los tiempos promedios para el diagnóstico, el suministro y el de trabajo y el número total de fallas detectadas en esos ítems, en el período observado.</p>
<p>Tiempo promedio para el diagnóstico</p>	$TPPD = \sum HTDMC / NTMC$	<p>Relación entre el tiempo en que se diagnostica una intervención correctiva en un conjunto de ítems con</p>

Capítulo 2. Propuesta de solución

		<p>falla y el número total de fallas detectadas en esos ítems, en el período observado. Utilizar en ítems para los cuales el tiempo de reparación o sustitución es significativo en relación al tiempo de operación.</p>
<p>Tiempo promedio para el suministro</p>	$TPPS = \Sigma HTSMC / NTMC$	<p>Relación entre el tiempo parado por piezas una intervención correctiva en un conjunto de ítems con falla y el número total de fallas detectadas en esos ítems, en el período observado. Utilizar en ítems para los cuales el tiempo de reparación o sustitución es significativo en relación al tiempo de operación.</p>
<p>Tiempo promedio para trabajo</p>	$TPPT = \Sigma HTTMC / NTMC$	<p>Relación entre el tiempo de trabajo de intervención correctiva en un conjunto de ítems con falla y el número total de fallas detectadas en esos ítems, en el período observado. Utilizar en ítems para los cuales el tiempo de reparación o sustitución es significativo en relación al tiempo de operación.</p>

Capítulo 2. Propuesta de solución

Tiempo promedio para fallar	$TPPF = \Sigma HROP / NTMC$	Relación entre el tiempo total de operación de un conjunto de ítems y el número total de fallas detectadas en esos ítems, en el período observado. Aplicar a ítems que se sustituyen después de fallar, sin reparar.
Utilización	$U = (HMR / HPM) * 100\%$	Utilización de equipos instalados. Horas de marcha reales (HMR) entre horas posible de marcha (HPM) por 100. Aplicable a centrales y equipos que se encuentran en funcionamiento y reserva.
Rendimiento	$R = (Producción Real / Período / Producción Nominal / Período) * 100$	Sintetiza el buen control del proceso productivo así como el estado de funcionamiento de las instalaciones.
Costo de MW generado	$CostMWG = Costo del MW / Generación$	Es el costo por cada Mega watts generado en un intervalo de tiempo.

Tabla 15. Indicadores Claves de Rendimiento.

2.7 Extracción, Transformación y Carga (ETL)

Los indicadores base obtenidos determinan la trayectoria a seguir por el proceso en el espacio de las configuraciones de las variables o dimensiones que determinan la dinámica de este. Siguiendo la metodología se le asignaron valores a los campos en cada tabla. Este proceso de

Capítulo 2. Propuesta de solución

asignación de valores se basa en una tecnología de **extracción** de datos del universo relacional de los OLTP, que conforman el entorno operacional de la planta, para, mediante un procedimiento de estandarización, limpieza y transformación, realizar el proceso de carga de los datos en cada entidad del almacén de datos. En la siguiente figura se puede observar la arquitectura básica del servicio de integración:

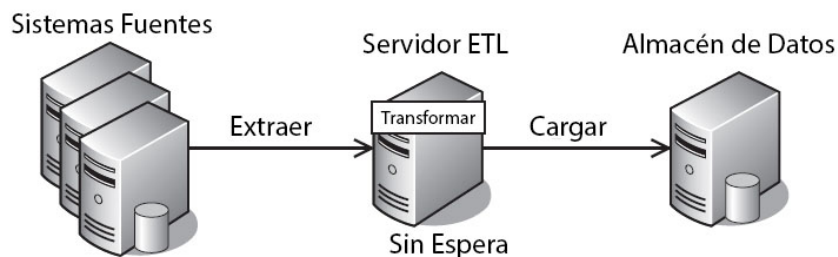


Figura 24. Arquitectura del Servicio de Integración.

Para ejecutar la carga del almacén de datos se diseñó un flujo de integración por cada entidad (ver Anexo 1), donde se realizan las operaciones de inserción, actualización y eliminación de los datos. En la siguiente figura se observa un ejemplo del flujo en el llenado de la dimensión Recursos Humanos (DimRrhh).

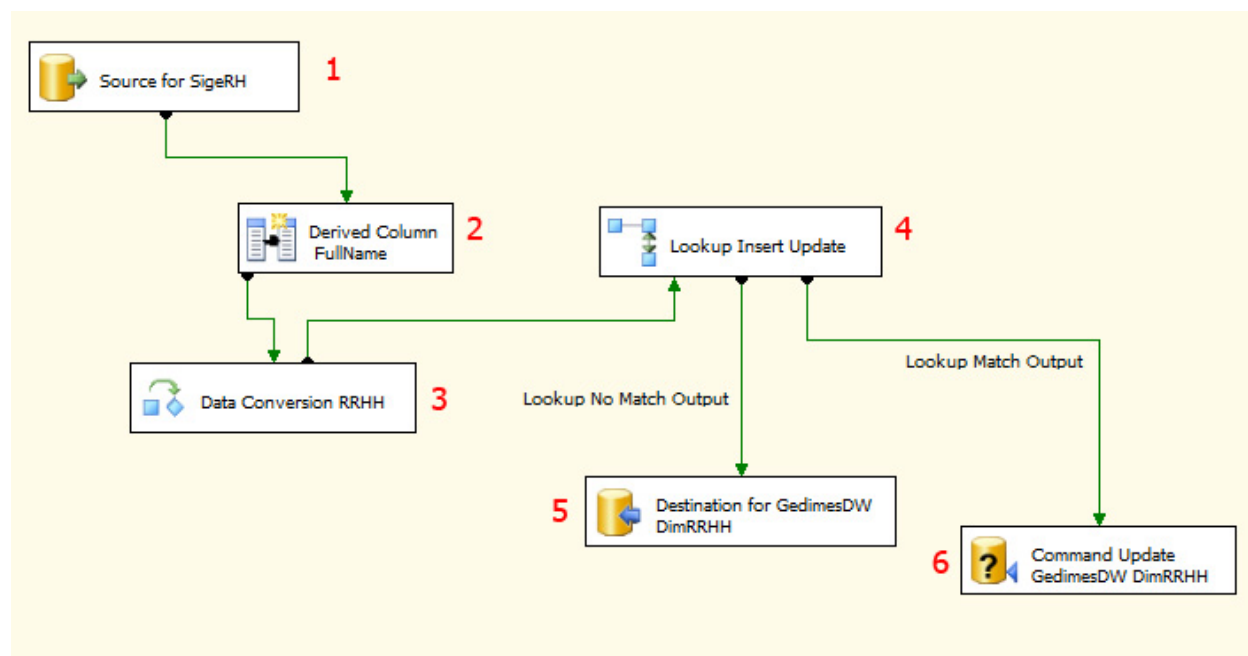


Figura 25. Diseño del flujo de integración para la dimensión RRHH.

El componente señalado con el número 1 es un objeto fuente transaccional donde se ejecuta una consulta transaccional SQL para obtener la tabla fuente con los campos requeridos hacia

Capítulo 2. Propuesta de solución

el componente de destino. El componente número 2 adiciona una columna donde se concatenará los campos (Nombre, PrimerApellido y SegundoApellido) para formar el campo NombreCompleto, a partir de esta operación en el componente 3 se convierten todos los campos del sistema fuente al mismo formato del destino. En el componente se busca el campo llave de la información transformada en el almacén de datos para definir cuál es la operación a realizar. Si no se encuentra el identificador la información es insertada en el componente 5. Si se encuentra se actualiza la información en el componente 6 del almacén de datos.

2.8 Implementación

Una vez implementado el servicio ETL sobre el Cubo MOLAP, se describirá la herramienta cliente que permite acceder a la tabla Pívor del Cubo, donde usuarios de la aplicación de Inteligencia de Negocio podrán conjugar las diferentes métricas con respecto a dimensiones que le pertenecen en el cubo, obteniendo estructuras vectoriales que brindan información del comportamiento retrospectivo y actual de las realizaciones dinámicas de las diferentes métricas del proceso.

2.8.1 Microsoft Excel 2010

Una parte fundamental de todo proyecto de DW/BI está en proporcionar a una gran comunidad de usuarios una forma más estructurada y por lo tanto, más fácil, de acceder al almacén de datos. Microsoft Excel 2010 se complementa perfectamente con el servicio de análisis implementado, ya que los proveedores de acceso a datos y las interfaces de interacción fueron diseñadas para trabajar en conjunto. Además los informes mensuales del proceso de generación de energía, deben ser enviados al SEN en un archivo de tipo Microsoft Excel, por lo que la utilización de esta herramienta facilita la confección e integración de los documentos en la organización.

Las operaciones realizadas con el cliente OLAP Microsoft Excel se muestran a continuación:

Capítulo 2. Propuesta de solución

1	Row Label	Energía Generada	Consumo Materia	Cant Arranque	Cant Agua Producida	Consumo Especifico de Combustible Status	Consumo Especifico de Combustible
2	Bueno	95290	8308	1706	8383	○	▲
3	Materia2					○	▲
4	Materia3	4921	504	112	378	●	▲
5	Materia7	2462	161	81	0	○	▲
6	Materia10					○	▲
7	Materia11	8665	556	183	587	○	▲
8	Materia14	2578	474	81	539	●	▲
9	Materia16	7375	996	117	654	●	▲
10	Materia24	76	344	14	340	●	▲
11	Materia26	4737	423	0	41	○	▲
12	Materia28					○	▲
13	Materia29					○	▲
14	Materia32	2011	673	168	668	●	▲
15	Materia33	7655	516	79	821	○	▲
16	Materia35					○	▲
17	Materia37	1278	364	75	13	●	▲
18	Materia38					○	▲
19	Materia42					○	▲
20	Materia44	3862	14	58	324	○	▲
21	Materia45	6925	620	174	469	○	▲
22	Materia46					○	▲
23	Materia49	3370	406	85	291	●	▲
24	Materia50	3996	320	83	329	○	▲
25	Materia63	4480	120	18	298	○	▲
26	Materia65	3859	333	10	210	○	▲
27	Materia67					○	▲
28	Materia75	4408	91	84	498	○	▲

Figura 26. Tabla Pívor Multidimensional usando Microsoft Excel 2010

Se cruzan las métricas del hecho Producción con la dimensión MateriaPieza permitiendo desagregar la información por los niveles de jerarquías especificados en la misma. Además se muestran los estados y tendencias de los KPI permitiendo el análisis del proceso de una manera dinámica y compleja. Los círculos rellenos en color negro, por ejemplo para una fija, significa que el consumo de la materia ubicada en la misma fila está en perfecto estado.

Otra opción de Microsoft Excel 2010 es la presentación de los resultados, a través de diferentes gráficos, ya sea en forma de columna, línea, pastel, (Ver Figuras. 24, 25, 26) etc. Ofreciendo una mayor posibilidad de comprensión e interpretación de la información para el apoyo a la toma de decisiones en la planta de generación de energía GEDICH.

Capítulo 2. Propuesta de solución

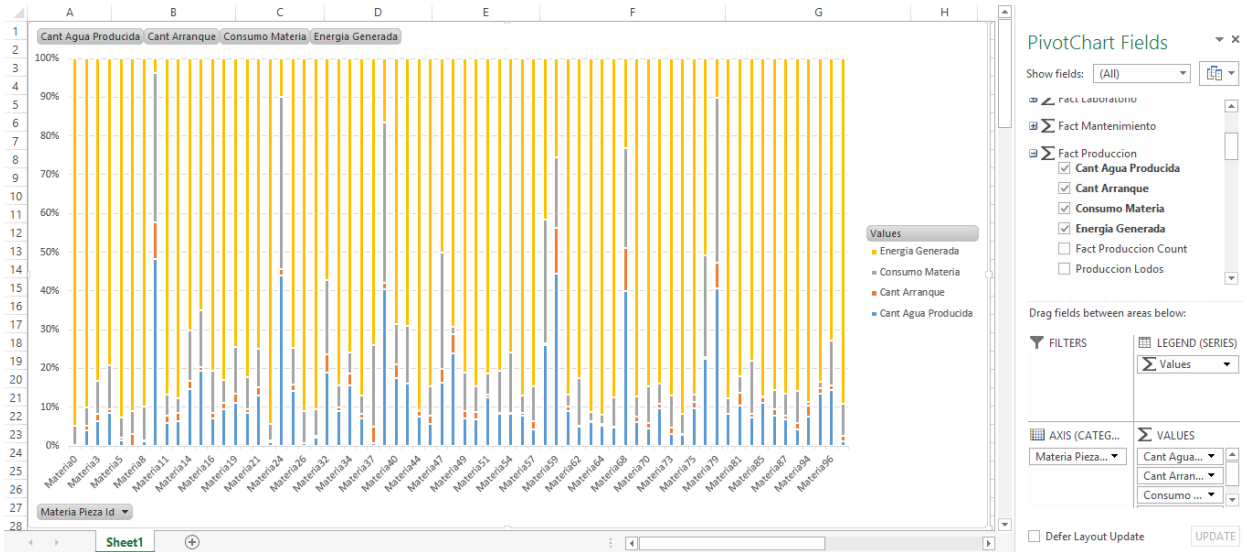


Figura 27. Gráfico Multidimensional en forma de columna usando Microsoft Excel 2010

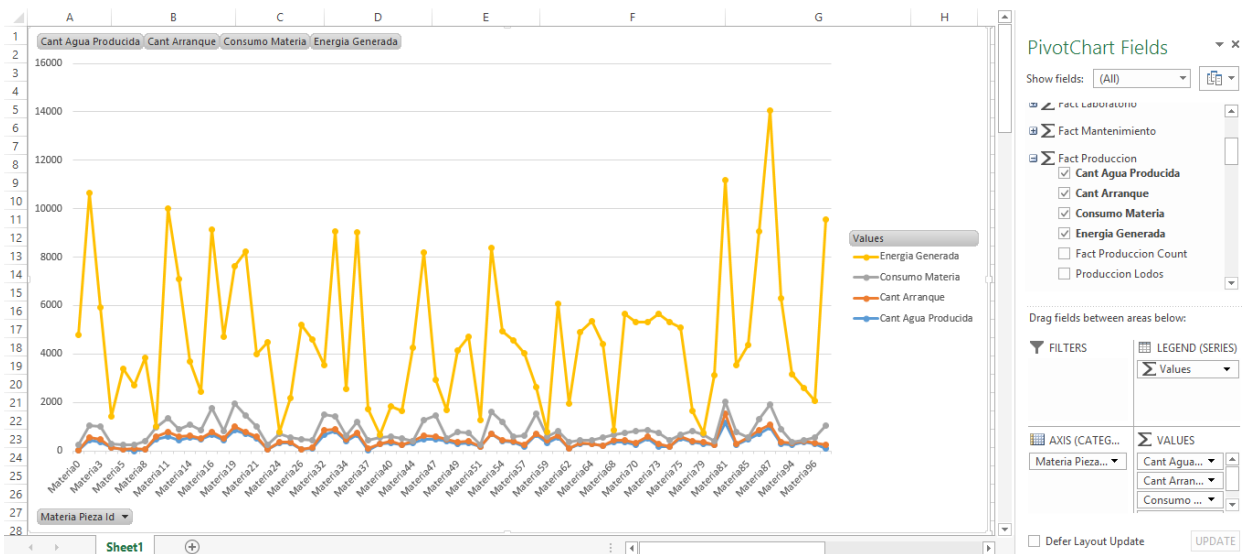


Figura 28. Gráfico Multidimensional en forma de línea usando Microsoft Excel 2010

Capítulo 2. Propuesta de solución

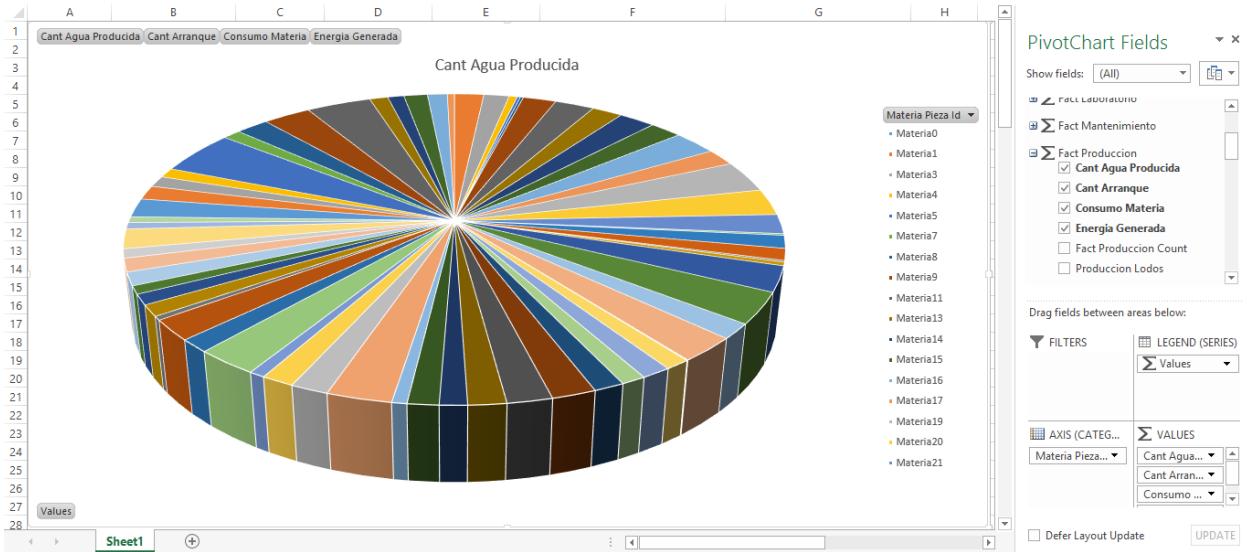


Figura 29. Gráfico Multidimensional en forma de pastel usando Microsoft Excel 2010

2.8.2 Microsoft SQL Server Agent (Agente SQL Server)

Para automatizar la ejecución de los paquetes de integración (ETL) soportados por SQL Server Integration Services se utilizó el Agente SQL Server. El Agente SQL Server es un servicio de Windows que permite automatizar tareas. Este debe estar activo para poder ejecutar automáticamente trabajos locales o multiservidor. Cada 10 minutos se ejecutan los paquetes configurados para así poder actualizar los datos en el almacén y así poder desplegar la aplicación del Cubo MOLAP.

Un ejemplo de la configuración del Agente SQL Server para la ejecución de un paquete de integración se muestra en la siguiente Figura:

Capítulo 2. Propuesta de solución

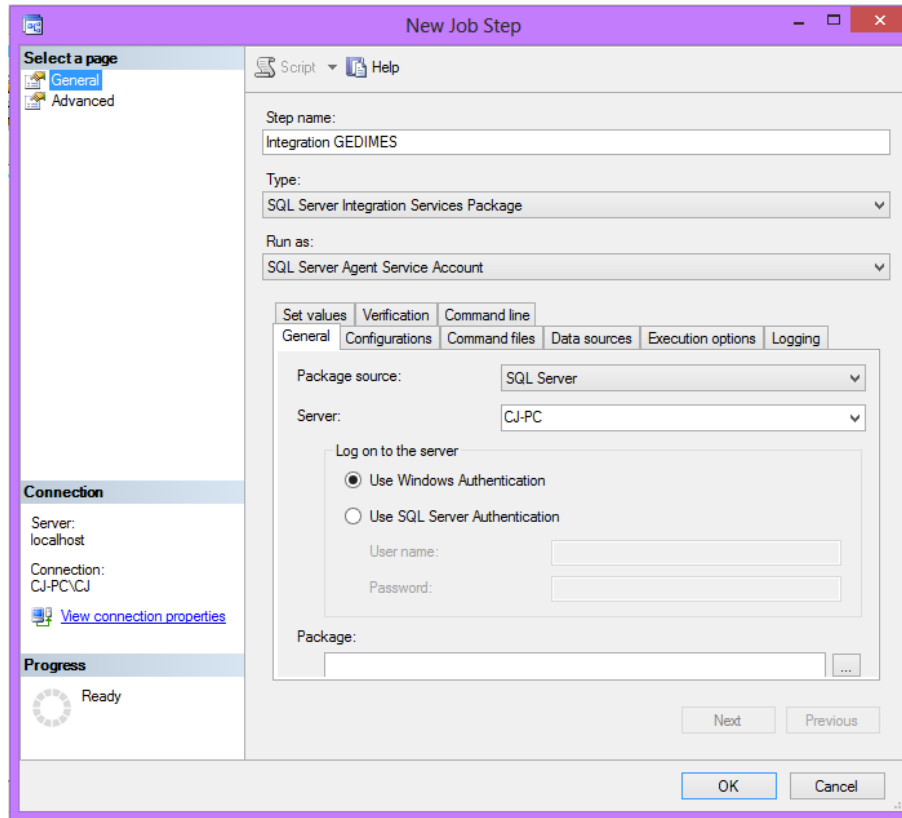


Figura 30. Configuración del Agente SQL Server para ejecutar un flujo de integración

2.9 Infraestructura de la solución de Inteligencia de Negocio.

La infraestructura para la solución de Inteligencia de Negocio propuesta, está conformada por la siguiente estructura de red, como se muestra a continuación:



Figura 31. Distribución genérica de la red para la Solución BI en GEDICH

Capítulo 2. Propuesta de solución

2.10 Conclusiones

En este capítulo se expusieron los resultados obtenidos para la construcción del DWH, siguiendo los pasos basados en la metodología de Kimball definida anteriormente. La especificación de requisitos del negocio y las restricciones del negocio constituyó la base para el modelado dimensional del Cubo OLAP. Se obtuvo la información necesaria para construir el almacén de datos y así proceder a la realización del proceso ETL, de vital importancia para el éxito de la solución.

3 Capítulo 3. Evaluación del análisis de datos

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 Generación de Energía Distribuida Ciudad de la Habana. 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.

3.1 Validación y pruebas

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. (10)
- 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. (11)

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 validar y verificar la calidad del sistema. Para el presente trabajo se tienen en cuenta los parámetros de escalabilidad, fiabilidad e integridad. Es por ello que se seleccionan los siguientes tipos de pruebas:

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. Se aplica para verificar el correcto funcionamiento del proceso ETL.

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, permitiéndole al administrador del sistema determinar si la aplicación rendirá lo suficiente en caso de que la carga real supere a la carga esperada. Se aplica para verificar la cantidad de usuarios conectados que soporta el sistema concurrentemente.

Capítulo 3. Evaluación del análisis de datos

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 del modelo: Si el modelo ya está estable, lo que se hace habitualmente es probarlo contra los requerimientos del negocio. Haciendo la pregunta práctica de ¿Cómo podemos obtener esta información en particular del modelo? Para las pruebas se pueden usar diseños de reportes estructurados, de usuarios actuales, diseños de cubos prospectivos, etc.

3.2 Resultados de la aplicación de las pruebas

3.2.1 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 12 procesos de ETL se tomó una muestra de 6 arrojando como resultado:

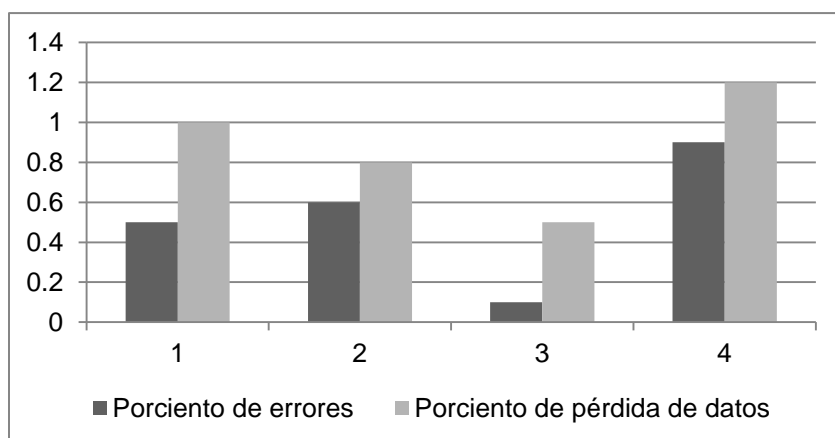
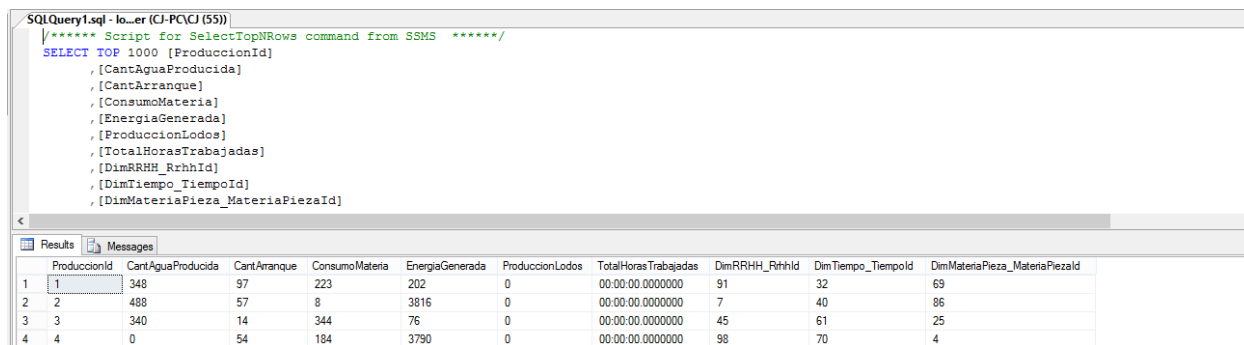


Figura 32. Por ciento de errores por iteración

A continuación se muestra un ejemplo de la prueba realizada a los datos del hecho producción:

Resultado de consulta realizada a la base de datos:

Capítulo 3. Evaluación del análisis de datos



```
SQLQuery1.sql - lo...er (CJ-PCVCJ (55))
/***** Script for SelectTopNRows command from SSMS *****/
SELECT TOP 1000 [ProduccionId]
, [CantAguaProducida]
, [CantArranque]
, [ConsumoMateria]
, [EnergiaGenerada]
, [ProduccionLodos]
, [TotalHorasTrabajadas]
, [DimRRRH_RrhhId]
, [DimTiempo_TiempoId]
, [DimMateriaPieza_MateriaPiezaId]
```

	ProduccionId	CantAguaProducida	CantArranque	ConsumoMateria	EnergiaGenerada	ProduccionLodos	TotalHorasTrabajadas	DimRRRH_RrhhId	DimTiempo_TiempoId	DimMateriaPieza_MateriaPiezaId
1	1	348	97	223	202	0	00:00:00.00000000	91	32	69
2	2	488	57	8	3816	0	00:00:00.00000000	7	40	86
3	3	340	14	344	76	0	00:00:00.00000000	45	61	25
4	4	0	54	184	3790	0	00:00:00.00000000	98	70	4

Figura 33. Datos que se encuentran persistidos en la base de datos

ProduccionId	CantAguaProducida	CantArranque	ConsumoMateria	EnergiaGenerada	ProduccionLodo	TotalHorasTrabajada	DimRRRH	DimTiempo	DimMateriaPieza
1	348	97	223	202	0	00:00.0	91	32	69
2	488	57	8	3816	0	00:00.0	7	40	86
3	340	14	344	76	0	00:00.0	45	61	25
4	0	54	184	3790	0	00:00.0	98	70	4

Figura 34. Datos consultados desde el cliente OLAP

Se puede concluir que se produce una pérdida de datos del 0.875% realizando un promedio de las cantidades de errores arrojados sobre el conjunto en general al realizarse el proceso de extracción, transformación y carga desde las fuentes de datos. Una pérdida de menos del 1% en la mayoría de las iteraciones se considera que no es significativa teniendo en cuenta la cantidad de datos que se consultan en cada extracción. Esta pérdida de datos está dada por la calidad de la información que se consulta desde los sistemas fuentes, aunque estos datos con errores se almacenan en archivos de texto persistentes en el servidor de integración, por lo cual pueden ser recuperados en su totalidad de forma manual. Mientras mayor sea la calidad de los datos, el porcentaje de perdidas será menor.

3.2.2 Resultados de la prueba de carga y estrés

Estas pruebas fueron realizadas sobre un ordenador con 4 Gb de RAM y un procesador AMD Phenom II X2 N660 3.0 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 200 usuarios, mostrándose errores en un 2% de las peticiones como se muestra en la siguiente figura:

Capítulo 3. Evaluación del análisis de datos

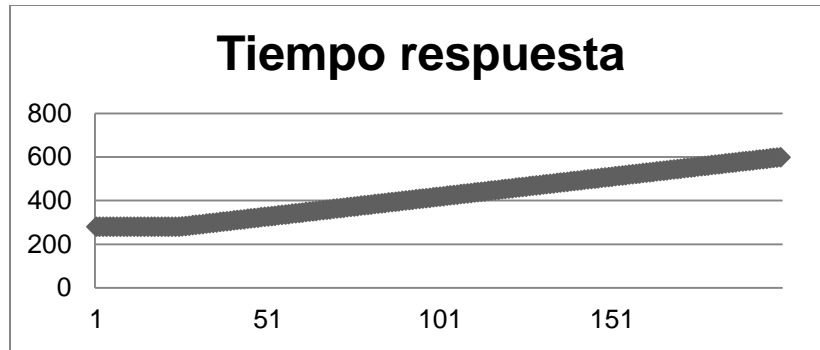


Figura 35. Muestra el tiempo medio de respuesta por la cantidad de usuarios

Se concluye que con las características de hardware relacionadas anteriormente la aplicación responde correctamente ante 190 peticiones concurrentes, lo cual supera la cantidad posible de usuarios que generalmente promedia los 40 usuarios. Es decir que soporta casi 5 veces la cantidad de usuarios promedio que acceden al sistema.

3.2.3 Resultados de la prueba del modelo

Para la realización de las pruebas del modelo fueron seleccionados expertos por cada área crítica del proceso de generación de energía distribuida para interactuar con la solución MOLAP de Gedimes. Cada experto generó los informes correspondientes a su área de trabajo, arrojando un resultado excelente esta prueba ya que se cumplieron todos los requisitos del negocio especificados.

3.2.4 Validación del Sistema

Una vez culminada la solución fue sometida a un proceso de revisión por los usuarios, culminando con la aceptación del sistema dada la importancia que tiene para dar soporte a la toma de decisiones de los grupos de dominio en el proceso de generación de energía distribuida proporcionando actualmente los siguientes beneficios:

- Recolectar y procesar la información generada a partir de los hechos críticos realizados en la planta.
- 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.

Capítulo 3. Evaluación del análisis de datos

En el Anexo 2 se muestra la evidencia del cierre del proceso de aceptación.

3.3 Conclusiones

Con la realización de las pruebas de integridad de datos, carga y estrés se comprobó el desempeño del servicio y los resultados fueron los esperados. La prueba de integridad de los datos mostrados en la realización de las tablas pivot, con respecto a los datos persistentes en la base de datos arrojó excelentes resultados, confirmándose que se produce una mínima pérdida de los datos al realizarse el proceso 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.

Conclusiones

Al culminar este trabajo se concluye que:

- La solución de Inteligencia de Negocio Gedimes BI sirve de base metodológica para el desarrollo, no solo de plantas de generación distribuida de energía eléctrica, sino para cualquier entidad manufacturera, ya que su diseño está basado en estándares internacionales.
- 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 tecnologías seleccionadas permitieron realizar una correcta elaboración de la solución a través de la explotación de las ventajas que estas ofrecen para la definición de los hechos y dimensiones que conforman la matriz bus, el diseño del modelo de datos para el almacén de datos, así como la arquitectura de la solución y la información.
- Los 12 procesos ETL diseñados permitieron realizar la extracción de los datos desde las fuentes, procesarlos y cargarlos hacia el almacén de datos actualmente en explotación.
- La implementación de la solución demostró ser el soporte idóneo en el nivel de gestión, para apoyar la toma de decisiones de los grupos del dominio en una planta de energía distribuida.

Recomendaciones

- Que se tome como referencia el contenido de esta investigación para próximos sistemas de inteligencia de negocio que se desarrollen.
- Desarrollar otras aplicaciones de inteligencia de negocio basadas en la analítica predictiva para poder profundizar en la toma de decisiones a largo plazo.

Referencias Bibliográficas

1. **Herrera, Ing. Silvio Dorta.** *MANUAL DE GESTIÓN (MAGEST)*. s.l. : Poligráfico Federico Engels, 2009.
2. Ecoticias. [En línea] 2008. <http://www.ecoticias.com/20080708-la-tecnologia-revolucionara-la-produccion-electrica-en-10-anos.html>.
3. Distributed Generation. [En línea] 2010. <http://www.distributed-generation.com/>.
4. **Langit, Lynn.** *Smart Business Intelligence Solutions*. s.l. : Microsoft Press, 2009.
5. **publications, IEC.** Enterprise control system integration 62264. s.l. : IEC publications, 2003. Vols. I, II, III.
6. **Inmon.** *Building the Data Warehouse*. s.l. : John Wiley & Sons, 2002.
7. **Kimball, Ross.** *The Data Warehouse Toolkit: The Complete*. s.l. : Wiley, 2002.
8. **Microsoft.** Libros en Pantalla SQL SERVER 2008 R2. s.l. : Microsoft, 2010.
9. **Rivadera, Gustavo R.** La metodología de Kimball para el diseño de almacenes de datos.
10. **Electrónicos, Instituto de Ingenieros Electricos y.** IEEE, std 610-1900.
11. **Pressman, Roger.** *Ingenieria del Software: Un Enfoque Practico*. s.l. : McGraw-Hill, 2006.
12. Sistemas de Apoyo a la Toma de Decisiones. [En línea] 2003. <http://www.datamart.el>.
13. Nexus Internet Solutions. [En línea] 2009. <http://www.nexus-soluciones.net/Artcle.aspx?aid=155>.
14. **Handbook, Comptroller's.** *Management Information Systems*. 1995.
15. **Guy Fouché, Lynn Langit.** *Foundations of SQL Server 2008 R2 Business Intelligence*. s.l. : Apress, 2008.
16. **Robert Wrembel, Christian Koncilian.** *Data Warehouses and OLAP*. s.l. : IRM Press, 2006.
17. **Rainardi, Vincent.** *Building a Data Warehouse*. s.l. : Apress, 2008.

Referencias bibliográficas

18. **Imhoff, Galemmo.** *Mastering Data Warehouse Design: Relational.* s.l. : Wiley Publishing, 2003.

19. **Kimball, Caserta.** *The Data Warehouse ETL Toolkit.* s.l. : Wiley, 2004.

20. *Desarrollo de una Plataforma para la Gestión de la Información en una Empresa de Generación Distribuida de Energía Eléctrica en Cuba (GEDIMES).* **Cobas, Luis Manuel Castillo.** La Habana : s.n., 2012.

Anexos

Anexo 1. Modelo de los flujos de Integración

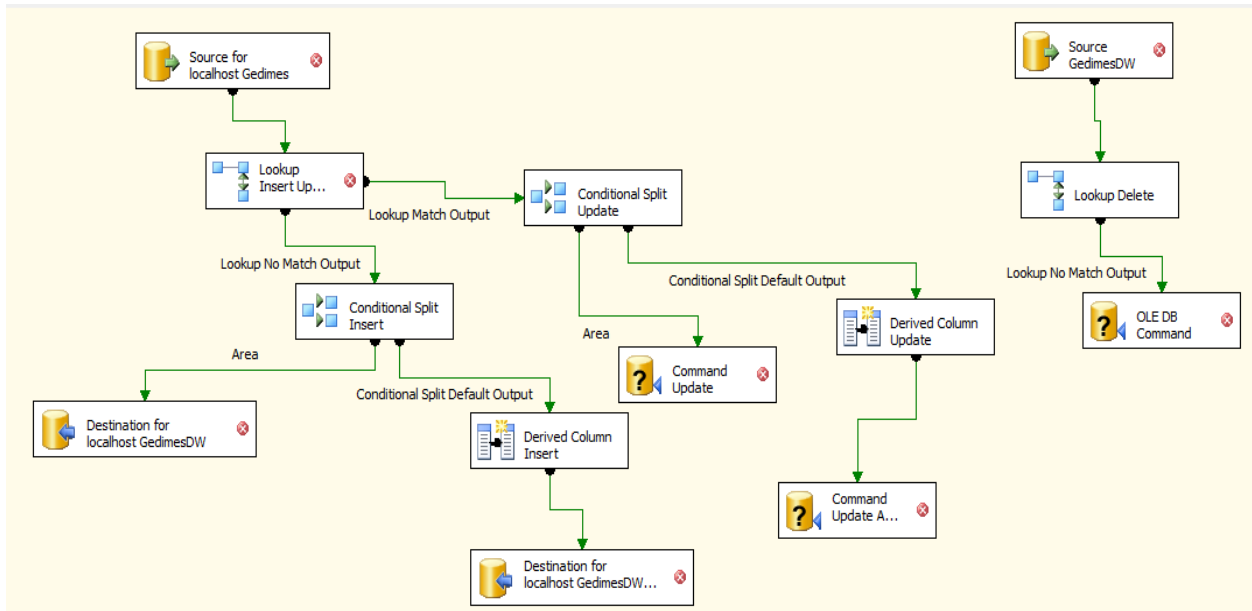


Figura 36. Diseño del flujo de integración para la dimensión RRHH.

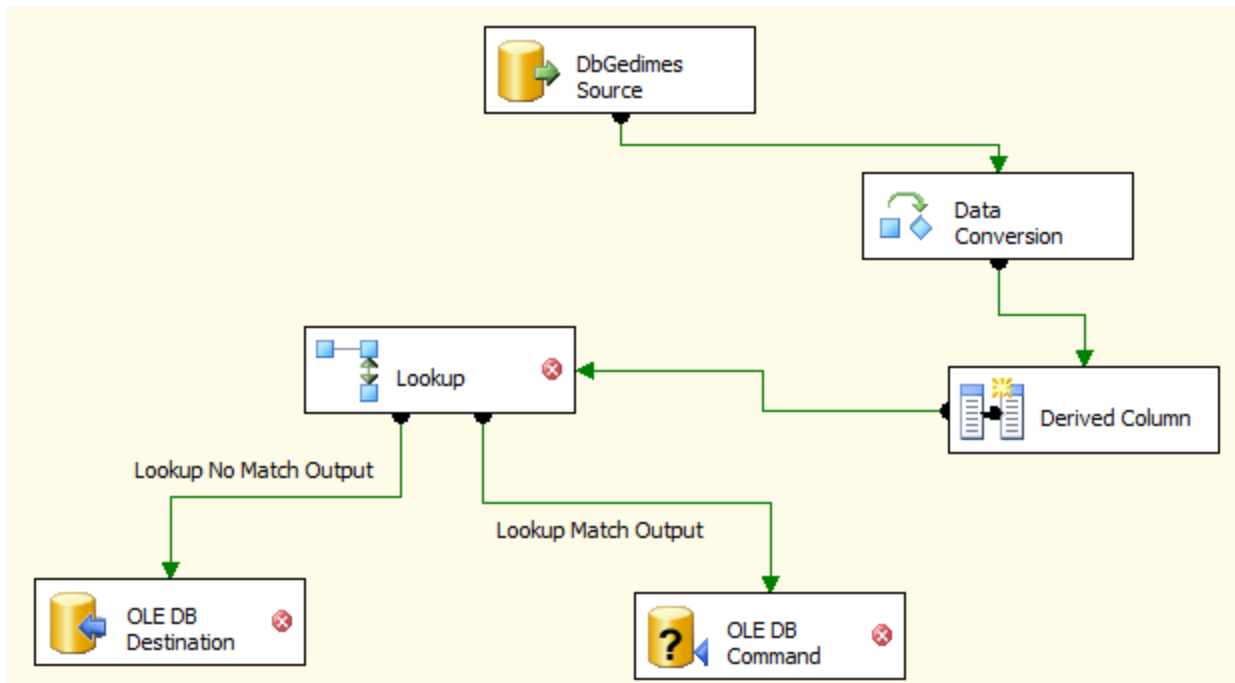


Figura 37. Diseño del flujo de integración para la dimensión Factura.

Anexo

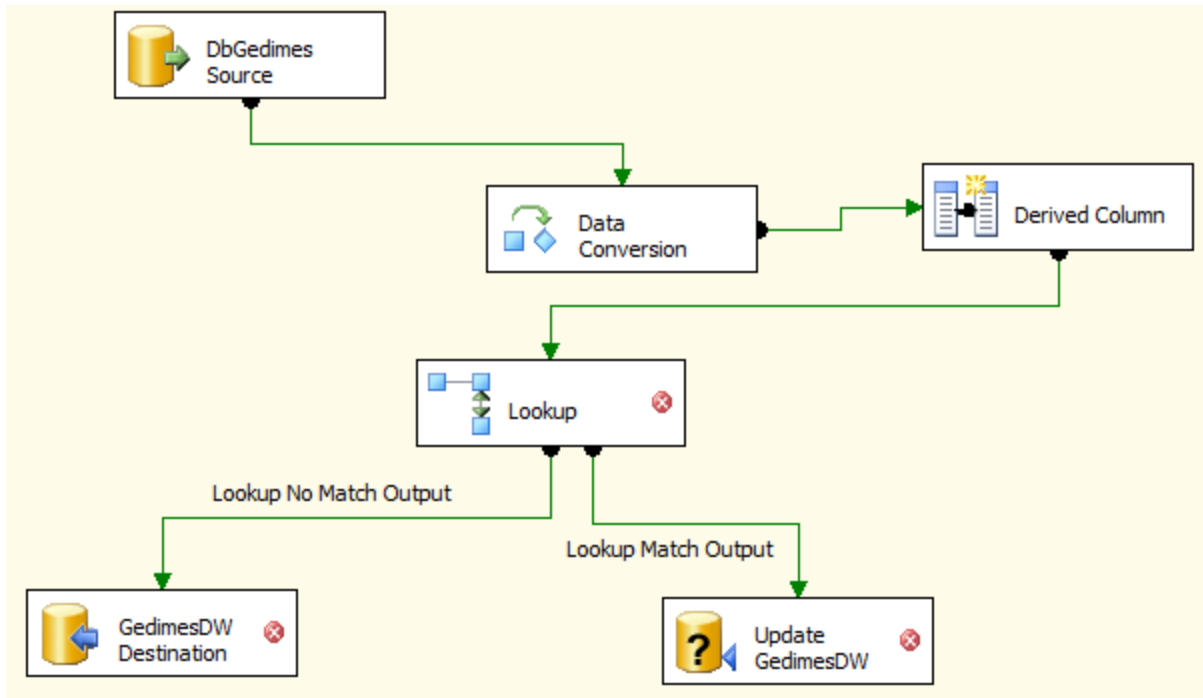


Figura 38. Diseño del flujo de integración para la dimensión Prueba.

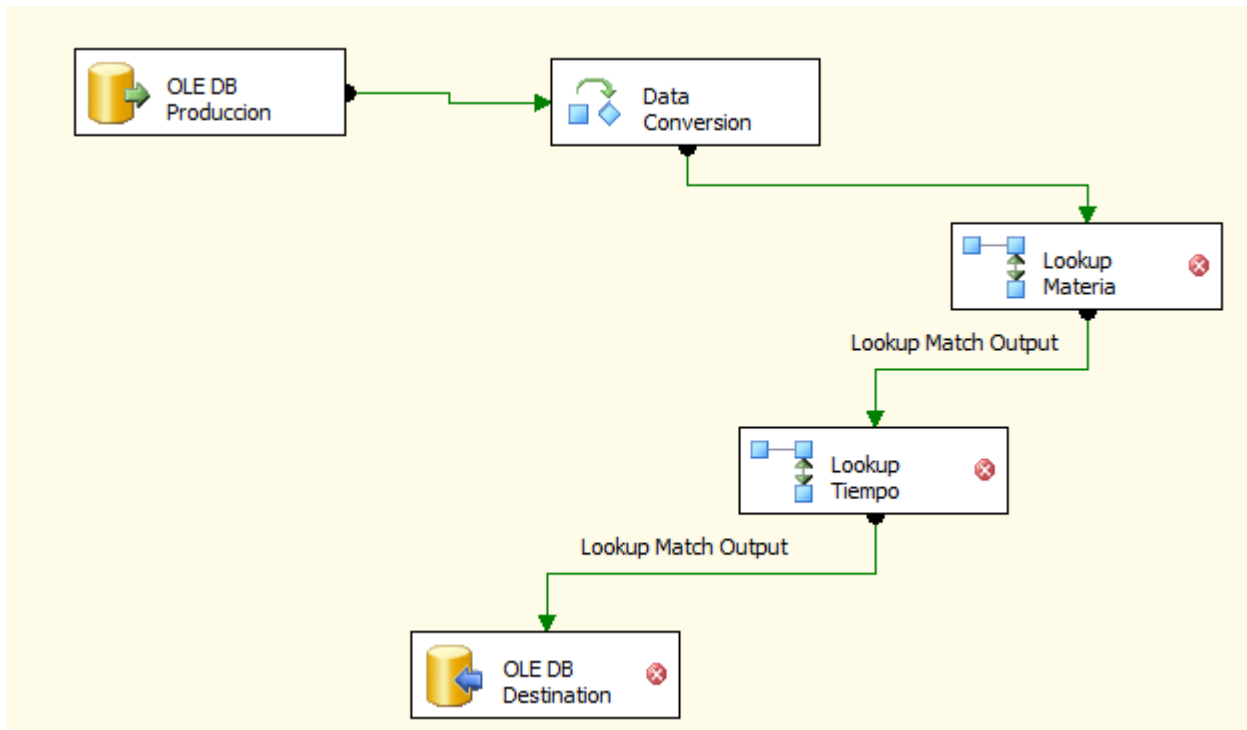


Figura 39. Diseño del flujo de integración para el hecho Producción.

Anexo

Anexo 2. Acta de aceptación



Empresa de Generación Distribuida Ciudad de la Habana.
Diaria No. 201 esq. Alambique, Habana Vieja.

La Habana, 3 de junio de 2013.
"Año 55 de la Revolución".

Acta de Aceptación

Proyecto: Desarrollo del módulo de Inteligencia de Negocio en un sistema de control de la producción para una planta de generación de energía distribuida

Producto: Gedimes BI

Categorías de las pruebas: Revisión a la aplicación

Fecha de conciliación: 25 de mayo del 2013

Observaciones del proceso:

Teniendo en cuenta que fue validada por el cliente la solución sin arrojar No Conformidades (NC), para la etapa de despliegue del sistema, se ha tomado el acuerdo de aceptar la solución de Inteligencia de Negocio (Gedimes BI) para un sistema de control de la producción en la planta de generación de energía distribuida de la Habana (GEDICH). Para que conste la aceptación de los resultados de las pruebas y por tanto la aceptación de los entregables especificados, dando fe al acuerdo, se extiende la presente Acta en tres (3) ejemplares, rubricados por los principales representantes de las partes.

Al tanto de cualquier aclaración en este importante asunto, le saludan fraternalmente

Ing. Carlos S. Rodríguez Silva
Esp. Principal Informática UEB-IAC

Ing. Alain Serrano Roque
Director de UEB- IAC de GEDICH



Figura 40. Acta de aceptación de la solución Gedimes BI