

# Arquitectura DATA WAREHOUSING para la operación de los Servicios de Tecnologías de la Información de ETECSA

DATA warehousing architecture for the operation of the information technologies services in ETECSA

MSc. Dariel Pérez Ferrer <sup>1</sup>, MSc. Alexander Gómez Betancourt <sup>2</sup>, MSc. Rafael E. Mazaira Fernández<sup>3</sup>, Dr.Sc. Lindsay Alonso Gómez Beltrán<sup>4</sup>

Recibido: 10/2018 | Aceptado: 03/2019

## **PALABRAS CLAVE**

Almacén de Datos;  
Inteligencia Negocio;  
Metodología.

## **RESUMEN**

En esta investigación se propone el desarrollo de una arquitectura Data Warehousing que permita monitorear los indicadores de la operación de los servicios de Tecnologías de la Información (TI) para mejorar el proceso de toma de decisiones. Por tanto, se caracteriza la gestión de TI para identificar indicadores y métricas que incidan directamente en la calidad del servicio. Además, se estudian diferentes enfoques, metodologías y herramientas BI vinculadas a los conceptos de almacén de datos y arquitectura Data Warehousing.

Como resultado de la investigación se diseñó e implementó la arquitectura propuesta, lo que permitió a los directivos de Tecnologías de la Información de ETECSA, obtener información integrada y veraz para la toma de decisiones en el proceso de gestión de TI.

## **KEYWORDS**

Data warehouse  
Business Intelligence  
methodology

## **ABSTRACT**

This research proposes the development of a Data Warehousing which allows to monitor indicators of Information Technology (IT) service operations to improve the decision making process. Therefore, IT management is characterized in order to identify indicators and metrics, which directly contribute to the quality of services. In addition, the author or authors analyses different approaches, methodologies and BI tools associated with data storage and data warehousing architecture.

As a result of this research, the proposed architecture was designed and implemented, enabling ETECSA's IT directors to obtain integrated and truthful information to make decisions during the IT management process.

1 Empresa de Telecomunicaciones de Cuba S.A., Cuba. dariel.perez@etecsa.cu

2 Empresa de Telecomunicaciones de Cuba S.A., Cuba. alexander.gomez@etecsa.cu

3 Empresa de Telecomunicaciones de Cuba S.A., Cuba. rafael.mazaira@etecsa.cu

4 Universidad de Camagüey "Ignacio Agramonte Loynaz", Cuba. lindsay.gomez.beltran@gmail.com

## Introducción

En la actualidad las empresas utilizan las Tecnologías de la Información (TI) para gestionar el negocio, por este motivo la tecnología ha pasado de ser un apoyo administrativo a constituirse en una herramienta de diferenciación y ventaja estratégica. Los volúmenes de información que requieren las organizaciones de hoy en día han ocasionado que se deba dar un trato especial a la gestión de las tecnologías de la información y las comunicaciones, por ello se han elaborado estándares y mejores prácticas tendientes a optimizar el uso de recursos tecnológicos, minimizar el impacto de los riesgos asociados y contribuir a la consecución de estrategias empresariales (Paredes, 2014).

Para un mercado tan competitivo como es el sector de las telecomunicaciones, Rodríguez y Almarcha (2006) plantean que los operadores necesitan un mayor grado de flexibilidad y la capacidad de reaccionar con rapidez ante nuevos requerimientos de las áreas de negocio. Estos objetivos son adicionales a los tradicionales de los responsables de sistemas de información, como la reducción de costos, el asegurar la disponibilidad de los servicios y garantizar que los sistemas den el rendimiento necesario. Como consecuencia de esta situación los departamentos de TI han pasado de ser considerados centros de coste, encargados de mantener los sistemas de información, a ser un verdadero socio estratégico para el área de negocio.

Por tanto, la actividad de TI como el resto de la organización necesita hacer uso de la información y transformarla en conocimiento, de manera que apoye la toma de decisiones, mejore la visión y permita definir estrategias. En este sentido, los almacenes de datos o Data Warehouse (DW) y las técnicas OLAP —*On-line Analytical Processing*— son las maneras más efectivas y tecnológicamente más avanzadas para integrar, transformar y combinar los datos que faciliten al usuario o a otros sistemas el análisis de la información (Pérez, 2015).

En este entorno la actividad de TI en ETECSA está caracterizada por la operación con sistemas de gestión que concentran los datos primarios para calcular indicadores, visualizar la información o mostrar estadísticas de los servicios que brinda en cada unidad organizativa. Sin embargo, esta información es extraída de manera manual, lo que hace que el trabajo se haga tedioso y lento, incurriendo en errores humanos e informaciones falseadas. Además, se evidencia la falta

de integración e interoperabilidad entre dichos sistemas pese a que los datos de estos, se encuentran almacenados en servidores de MySQL o PostgreSQL, como son los casos de GLPI y DemandasTI respectivamente, o pueden ser consultados a través de servicios web y relacionados por elementos comunes como: usuarios, entidades, equipamientos, tipos, entre otros.

A fines del primer semestre del 2016, en el sistema GLPI encargado de gestionar los activos informáticos, las interrupciones y las solicitudes de servicios de TI en la Empresa se comportaban de la siguiente manera: 15 154 computadoras, 11 902 monitores, 6 440 impresoras, 2491 elementos del red, 28 643 dispositivos y 12 667 usuarios, con una media mensual de 5 800 tickets. Sin embargo, el análisis de la información generada en este proceso se limita al chequeo de incidencias o solicitudes de servicios abiertas y cerradas, sin medir a fondo el trabajo de las mesas de ayuda o el personal de soporte que posibilite identificar las áreas de mejora en la actividad.

En teoría, las tecnologías deberían aportar mayor sistematización y capacidad de control, no obstante, la realidad dista mucho de esta visión. La actividad de TI en la Empresa no tiene un sistema robusto que soporte la toma de decisiones. El excesivo volumen de datos, la duplicidad de estos y la poca síntesis de la información, dificulta el trabajo de los directivos. Cuando se emplean sistemas informáticos destinados a garantizar los servicios de TI con fines estadísticos, se afecta su rendimiento. La falta de análisis de causas, los comportamientos y las tendencias de interrupciones en el tiempo, aportan una visión incompleta del negocio y como consecuencia limita la toma de decisiones.

Como consecuencia de lo anterior en esta investigación se desarrolla una arquitectura Data Warehousing que permita monitorear los indicadores de la operación de los servicios de TI para mejorar el proceso de toma de decisiones en ETECSA.

## Contenido

### Gestión TI

Las organizaciones deben ir de la mano con la tecnología, pues es una ventaja competitiva ya definida (Porter, 2015); por tanto, invierten en la adquisición de hardware y software con la concepción de que se trata de una importante inversión, la debilidad entonces se traslada en su mayor parte a la gestión que dan a estos.

Según Mabilon (2010) gestionar pasa por establecer objetivos y prioridades, valorar mejores opciones, asignar recursos, gestionar y evaluar la ejecución, orientando y alineando la mejora del negocio. Según su criterio, la gestión de TI abarca cuatro funciones claves:

- Gestión de la Arquitectura TI
- Planificación y Gobierno TI
- Gestión del ciclo de vida TI
- Gobierno del aprovisionamiento

### Marcos de trabajo y estándares internacionales

Varios investigadores han discutido la necesidad de gestionar TI soportado en estándares y buenas prácticas (Piattini y Hervada, 2007; Sunthonwutinin y Chooprayoon, 2013) para así obtener su mejor rendimiento, destacándose entre otros las siguientes normas o marcos de referencias:

- ISO38500, COBIT y Val IT: estándar para el gobierno corporativo de las TI y marcos de trabajo para el Gobierno TI
- PMBOK y Prince: metodologías y mejores prácticas para la gestión de proyectos.
- ISO/IEC 20000 y ITIL V3: prácticas de Gestión de Servicios TI
- ISO/IEC 27002: código de buenas prácticas de seguridad.
- ISO/IEC 27001:2005: especificaciones para los Sistemas de Gestión de la Seguridad de la Información (SGSI).
- ISO/IEC 24762:2008: guías de Seguridad para la continuidad de negocio.

De forma similar, otros autores como Calderon y Pompa (2012) y Mujica y Arzola (2007) destacan la importancia de eTOM —*enhanced Telecommunication Operations Map*— como marco referencial de organizaciones de telecomunicaciones que orienta con las mejores prácticas el desarrollo y gestión de procesos, lo que permite identificar cuáles procesos e interfaces son más factibles de integrar, automatizar y estandarizar.

Sin embargo, para lograr una gestión efectiva de las TI no es obligatorio seguir un único marco de referencia, los autores Cocanguilla y Santino (2016) y Falcony (2013) integran en sus investigaciones las mejores prácticas de eTOM, ITIL —*Information Technology Infrastructure Library*— y Cobit. Anteriormente, Mujica y Arzola (2008) analiza com-

parativamente los marcos referenciales eTOM e ITIL con la norma ISO 9001:2000 y establece una relación entre ellos desde el punto de vista de su interacción en las organizaciones a fin de garantizar su eficacia, eficiencia y calidad al momento de implementarlos.

En síntesis, todos estos marcos referenciales, modelos y normas se han desarrollado con un mismo propósito: mejorar la eficiencia, facilitar la calidad de la gestión, incrementar la efectividad de los costos, reducir riesgos, entre otros.

### Inteligencia de Negocios

Howard Dresden, analista de Gartner, en 1989 propone una definición formal de Inteligencia de Negocios o Business intelligence (BI): “conceptos y métodos para mejorar las decisiones de negocio mediante el uso de sistemas de soporte basados en hechos”. Desde entonces, el concepto ha evolucionado (Curto, 2010).

Para Davenport y Harris (2007) la Inteligencia de Negocios consiste en la obtención, administración y reporte de la data orientada a la toma de decisiones, y las técnicas analíticas y procesos computarizados que se usan para el análisis de la misma.

Cuando se habla de Inteligencia de Negocios hay que considerar los diferentes elementos que la constituyen, dentro de los cuales están: la base de datos centralizada —*Data Warehouse*—, el conjunto de herramientas que utilizará el usuario final —*Business Analytics*—, las relaciones no conocidas entre las variables, que tienen que descubrirse mediante la minería de datos (también minería de texto y de la web), y metodologías complementarias como BPM —*Business Performance Management*—, las cuales sirven para monitorizar el desempeño y obtener ventaja competitiva (Turban, Volonino, y Pollard, 2008).

Diferentes autores (Curto, 2010; Inmon, 2005; Kimball y Ross, 2013) coinciden en que el componente principal de una solución de Inteligencia de Negocios es el almacén de datos. En este sentido, Trujillo y Song (2008) definen como Data Warehousing al proceso por el cual las empresas extraen sentido y significado de sus datos, a través del uso de un repositorio de datos o Data Warehouse.

Según Kimball y Ross (2013) la estructura básica de la Arquitectura Data Warehousing incluye:

- Datos operacionales
- Extracción de datos
- Transformación de datos

Carga de datos

Data warehouse

Herramientas de acceso al componente físico DW

Por lo antes mencionado, se considera importante para el desarrollo de la investigación profundizar en los elementos estructurales de la arquitectura *Data Warehousing*, en los acápites siguientes se caracterizan los aspectos teóricos sobre DW, arquitecturas, metodologías de diseño, ETL y herramientas de acceso a los datos.

## Data Warehouse

Un Data Warehouse según Inmon (2005), 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, almacenado en una base de datos diseñada para favorecer el análisis y la divulgación eficiente de estos.

Mientras, Kimball, Ross, y others (2002) definen Data Warehouse como: “una copia de los datos transaccionales estructurados específicamente para consultas y análisis”.

Dentro de un DW existen dos tecnologías que se pueden ver como complementarias, una relacional para consultas y una multidimensional para análisis. Así un DW resulta ser un recipiente de datos transaccionales para proporcionar consultas operativas, y la información con el fin de llevar a cabo el análisis multidimensional (Kimball y Ross, 2013).

Sin embargo, el acceso a los datos de toda la empresa a veces no es conveniente (o necesario) para determinados usuarios que solo necesitan un subconjunto de estos, en estos casos se utilizan los Data Marts. El concepto Data Mart es una especialización del DW, y está enfocado a un departamento o área específica, como por ejemplo los departamentos de finanzas o marketing, lo que permite un mejor control de la información que se está abarcando (Kimball y Ross, 2013).

## Arquitecturas

Las corrientes surgidas sobre Data Warehousing tienen varios enfoques de diseño e implementación para la gestión de bases de datos gerenciales, cada una con sus propias características y ventajas, a saber (Aranibar, 2013):

Top-Down: el proceso ETL concentra la información proveniente de las fuentes transaccionales en una sola base de datos gerencial corporativa a partir de donde se pueden generar múltiples Data Marts.

Bottom-Up: el proceso ETL genera Data Marts independientes que luego se consolidan en un Data Warehouse corporativo.

Hybrid: rescata las fortalezas de los enfoques top-down y bottom-up para crear un escenario de coexistencia entre ambos.

Federated: enfoque poco aplicado, pero habilitado para cambios permanentes de las exigencias de los usuarios de bases de datos gerenciales.

El modelo de Kimball (bottom-up) se adapta mejor a proyectos que se van a realizar por partes (Data Marts), con equipos más generales y pequeños, tomando los expertos el manejo de las métricas, extracción, limpieza, transformación y carga de datos.

Teniendo en cuenta todo lo analizado hasta este punto se decide emplear el enfoque de construcción bottom-up de Kimball debido a las características y el alcance departamental de esta investigación, acotada a la actividad de soporte TI de ETECSA.

## Metodologías de diseño

Un paso lógico en el desarrollo de cualquier arquitectura es la elección de la metodología a utilizar.

Hay muchas metodologías de desarrollo que pueden ser utilizadas para el desarrollo de un DW, entre ellas se pueden mencionar:

Oracle —*Harmony Methodology*—

SAS —*Rapid Data Warehouse Methodology*—

Kimball —*Ciclo de Vida. Dimensional del Negocio*—

Hefesto

En su investigación Garcés (2015) establece una comparación entre 18 metodologías, agrupadas en 4 categorías: tradicionales, BI genéricas, BI propietarias y ágiles. Para esto utiliza 68 criterios, que cubren las fases siguientes: requerimientos, estrategia de proyecto, planificación del proyecto, selección de la tecnología, diseño del sistema de información, elaboración del sistema de información, planificación de la implantación, implantación piloto, capacitación y puesta en marcha del sistema. Al analizar las calificaciones de las metodologías, Kimball y Hefesto obtienen los mejores resultados cuando se analizan en conjunto y tiene un peso significativo el tamaño, tiempo y costo de la implantación.

Por tanto, se selecciona la metodología Hefesto cuya ventaja principal (Brizuela y Blanco, 2013) es que especifica puntualmente los pasos a seguir en cada fase a diferencia de otras metodologías que mencionan los procesos, mas no explican cómo realizarlo.

### Diseño multidimensional

Las metodologías de diseño de DW de forma general definen varias tareas, entre las cuales se encuentra el diseño de la data que incluye el modelamiento y la normalización.

Un modelo multidimensional está formado de una tabla con llave compuesta, llamada tabla de hecho, y un conjunto de dimensiones que tienen una sola llave primaria que corresponde exactamente a una parte de la llave compuesta de la tabla de hecho.

La colección de tablas de hechos y dimensiones se conoce como esquema multidimensional. En (Inmon, 2005; Kimball, Reeves, Ross, y Thornthwaite, 1998; Peralta, 2001) se identifican los posibles esquemas. En (Bernabeu, 2010; Peralta, 2001) se abordan tres de los fundamentales:

Esquema en estrella: conformado por una tabla de hecho con un grupo de dimensiones

Esquema copo de nieve: es una variación del esquema de estrella en las que las tablas de dimensiones se organizan jerárquicamente

Esquema constelación de hecho: constituido por un conjunto de tablas de hechos que comparten algunas dimensiones

### Extracción, transformación y carga

Existe un consenso entre los autores (Jörg y Dessloch, 2009; Kimball y Caserta, 2004) que identifican a los procesos Extracción, Transformación y Carga (ETL) como el centro de cualquier DW.

Según Montalvillo (2012) este proceso consume entre 60% y 80% del tiempo de un proyecto de Inteligencia de Negocios y un referente como Kimball y Caserta (2004) señala que 70% de los recursos necesarios para la implementación y mantenimiento de un DW son típicamente consumidos por procesos ETL.

Durante los procesos ETL, los datos son extraídos de múltiples y heterogéneas fuentes como bases de datos de Procesamiento de Transacciones en Línea, archivos de texto, hojas de cálculo, entre otras. Luego son transformados y limpiados con el objetivo de ajustarlos al diseño del DW y cargados en el mismo don-

de son accesibles para las aplicaciones de Inteligencia de Negocios (Berson y Smith, 1997; El-Sappagh, Hendawi, y El Bastawissy, 2011).

En tanto Kimball, Ross, Thornthwaite, Mundy y Becker (2008) se definen 34 subsistemas para organizar las diversas tareas que son parte de un proceso ETL. Estos subsistemas se agrupan en:

Extracción (subsistemas del 1 al 3): obtención de los datos de sus fuentes de origen

Limpieza y conformación (subsistemas del 4 al 8): consiste en acciones que permiten validar y aumentar la calidad de la información

Entrega (subsistemas del 9 al 21): carga y actualización de los datos del Data Warehouse

Gestión (subsistemas del 22 al 34): controla y supervisa la correcta tramitación de todos los componentes de la solución ETL

### Herramientas BI

Las herramientas de Inteligencia de Negocio son aplicaciones digitales diseñadas para colaborar con el BI durante el análisis y la presentación de los datos.

Tomando en cuenta un estudio realizado por la empresa española Stratebi (2017) donde se presenta, analiza y compara 7 de las herramientas más potentes del entorno Business Intelligence: Tableau, PowerBI, Information Builders, SAS, QlikView, Amazon QuickSight y Pentaho se aprecia que las herramientas mejor posicionadas son SAS y Pentaho.

Sin embargo, una comparación previa realizada por Garcés (2015) donde analiza 20 herramientas BI en el periodo del 2010 al 2014, excluye a SAS de las 5 primeras herramientas BI propietarias, pero mantiene a Pentaho como segunda opción Open Source. Para esta investigación la autora analiza 72 criterios de evaluación tomados de los reportes anuales de cuatro organizaciones especialistas en la materia: Gartner, Forrester, BI Survey BARC —*Business Application Research Center*— y Dresner Advisory Services.

Por tanto, se selecciona la Suite Pentaho como herramienta BI para el desarrollo y despliegue de la arquitectura propuesta en esta investigación, al ser una plataforma desarrollada en Java bajo la filosofía de software libre, que incluye soluciones integradas para generar informes parametrizables, minería de datos, ETL, análisis multidimensional OLAP, análisis Ad Hoc, generación de alertas, entre otros.

En resumen, hasta este punto se han revisado los referentes teóricos asociados a los componentes de una arquitectura Data Warehousing, permitiendo seleccionar el enfoque, la metodología y la herramienta BI a utilizar en esta investigación.

### **Análisis de la información en la operación de los servicios de TI en ETECSA**

La gestión de las TI en ETECSA, se rige por procesos, legislaciones y procedimientos que se apoyan en normas y estándares internacionales como ITIL, la ISO 9000:2000 e ISO 9001:2000, entre otros y en resoluciones internas o del Ministerio de la Informática y Comunicaciones de Cuba como la 127/2007.

En el procedimiento PE-TI-037/10 se define la forma de evaluación, el análisis, el formato de entrega y el control de los indicadores, para medir el uso y la disponibilidad de los recursos tecnológicos, así como el impacto económico de las TI en ETECSA.

En el procedimiento antes mencionado se definen 19 indicadores, de ellos 5 están relacionados directamente con la operación de los servicios TI:

Disponibilidad de los servicios TI

Tiempo promedio de solución

Tiempo medio de atención a incidencias

Tiempo medio de ABM usuarios

Efectividad en el cumplimiento del plan de mantenimiento

Para el cálculo de estos 5 indicadores se utiliza la información proveniente de tres sistemas Nagios, GLPI y SAP-PM, de los cuales se extraen manualmente y se llenan los modelos establecidos en el procedimiento.

Sin embargo, el análisis de los indicadores por sí solo, aunque puede dar una idea de cómo marcha la actividad su valor es incompleto y los directivos encargados de la toma de decisiones tienen una visión limitada, pues no se profundiza en las causas, no se analizan las tendencias o comportamiento en el tiempo, sin incluir el análisis desde distintas perspectivas.

Además, el empleo de sistemas operacionales para conformar los indicadores de una forma manual es un proceso lento que puede llevar a errores. Asimismo, la necesidad de información que no esté comprendida en reportes de la aplicación conlleva a consultas directas sobre las bases de datos que pueden afectar el rendimiento de los aplicativos y ocasionar afectación del servicio.

El estudio del marco de referencia ITIL como parte de esta investigación aporta nuevos elementos a

tener en cuenta para medir la calidad y el trabajo de las mesas de ayuda, como la cantidad de incidencias atendidas en 1 nivel y 2 nivel, la cantidad y tiempo de solución de incidencias graves, entre otros, que hoy no son analizados.

En conclusión, aplicaciones como GLPI se centran en la gestión de incidencias o solicitudes y en el control de la técnica instalada pero su enfoque es reactivo y en consecuencia necesita ser apoyado por sistemas preventivos y predictivos, que permitan el análisis de los datos en el tiempo para encontrar posibles mejoras o eliminar causas de problemas. Por tanto, la Empresa requiere de mecanismos más eficientes que posibiliten monitorear la operación de los servicios de TI y apoyen la toma de decisiones por parte de los directivos.

### **Arquitectura propuesta**

Los estudios realizados permitieron establecer los elementos esenciales que forman parte de la arquitectura Data Warehousing propuesta. (Figura 1)

El diseño de esta arquitectura sigue el enfoque de construcción bottom-up donde el origen de los datos será el sistema GLPI, aplicación web desarrollada en PHP con un gestor de base datos Mysql Server.

Para el proceso ETL se seleccionó la herramienta Pentaho Data Integration, perteneciente a la Suite Pentaho mediante la cual se transforman e integran datos entre sistemas de información existentes y los Data Marts que compondrán el sistema BI.

Como gestor de base de datos, para el diseño, desarrollo e implementación del almacén de datos se empleará un servidor PostgreSQL v 9.3, debido fundamentalmente a la estabilidad, potencia, robustez, facilidad de administración e implementación de estándares de este gestor de Base de Datos.

El procesamiento analítico en línea OLAP, es el componente más poderoso del Data Warehousing (Bernabeu, 2010). En la arquitectura propuesta Pentaho Analysis Service fue seleccionado como servidor OLAP al estar integrado en la Suite BI de Pentaho.

Como elementos finales de la arquitectura se encuentran las herramientas que permitan realizar análisis, confeccionar reportes y cuadros de mandos. Pentaho Reporting, Community Dashboard Editor (CDE) y las vistas de análisis (jPivot y Saiku), son los encargados de garantizar esta actividad.

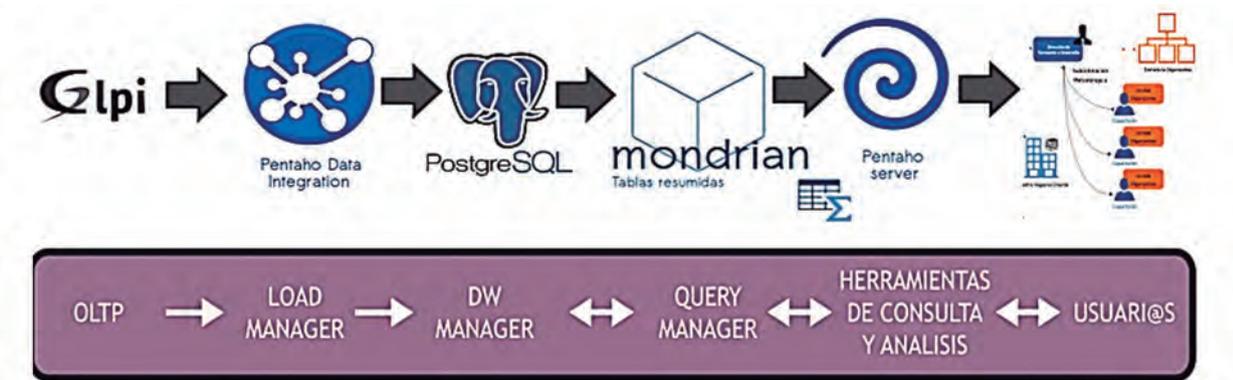


Figura 1. Arquitectura propuesta

## 2.6. Aplicación de la metodología

A continuación se desarrollan las cuatro fases definidas en la metodología Hefesto, seleccionada para el desarrollo de la investigación y a las características propias de la metodología definidas por Bernabeu (2010)

### Análisis de requerimientos

En esta fase se identificaron los requerimientos de los usuarios a través de preguntas que expliciten los objetivos de la organización. Luego, se analizaron estas preguntas con el fin de identificar cuáles fueron los indicadores y perspectivas que se tuvieron en cuenta para la construcción del DW. Finalmente, se confeccionó un modelo conceptual que permitirá visualizar el resultado obtenido en este primer paso.

Se conformó un equipo de trabajo integrado por directivos de la División de Tecnologías de la Información, jefes de departamento de TI en divisiones territoriales y especialistas de la actividad. Se establecieron encuentros de trabajo donde se analizaron los indicadores asociados a la operación de los servicios TI definidos en el procedimiento PE-TI-037/10 del Proceso TI. Además, se tomaron como guía las métricas operativas y los KPIs relacionados con la gestión de incidencias en el marco de trabajo ITIL.

Se les preguntó cuáles eran, según ellos, los indicadores que representan de mejor modo la operación de los servicios de TI y qué es exactamente lo que se desea analizar de este. La respuesta obtenida remitió a tener en cuenta y consultar datos sobre la cantidad de tickets resueltos, abiertos y pendientes, además de los tiempos de resolución, en espera y atención de incidencias o solicitudes.

Luego de analizar las preguntas del negocio como se indica en la metodología, se obtienen los siguientes indicadores: tickets cerrados, tiempo medio de solu-

ción, tiempo medio en espera, tiempo medio de atención, tiempo medio de afectación

Con sus respectivas perspectivas de análisis: nivel resolución, solicitante, técnico asignado, abierta por, tipo, entidad, tipo solución, categoría, nivel resolución, origen incidencia, prioridad, estado, fecha creado, fecha solución e incidencia.

### Análisis de los OLTP

Seguidamente, se analizan las fuentes OLTP para determinar cómo son calculados los indicadores y para establecer las respectivas correspondencias entre el modelo conceptual creado en el paso anterior y las fuentes de datos (base de dos mysql del sistema GLPI). Luego, se define qué campos se incluyen en cada perspectiva. Finalmente, se ampliará el modelo conceptual con la información obtenida en este paso.

### Modelo lógico del Data Warehouse

En este punto, se confecciona el modelo lógico de la estructura del almacén de datos, teniendo como base el modelo conceptual que ya ha sido creado. Tomando en cuenta las características del proyecto y lo estudiado se seleccionó el esquema en estrella, porque su uso permite optimizar el tiempo de respuesta de los reportes en comparación al modelo copo de nieve donde resulta más complejo.

Para definir las tablas de dimensiones se analizaron las perspectivas obtenidas en los pasos anteriores; se creó un nombre para cada tabla dimensión con una llave primaria; y además se renombraron los atributos que se consideraron necesarios, en total se crearon 14 dimensiones.

De igual forma, se creó una tabla de hecho que incluye las medidas (indicadores obtenidos) como parte de sus atributos y las claves foráneas de las dimensiones.

Finalmente, en esta fase se obtiene el esquema lógico del DW, al establecer las uniones entre las tablas de dimensiones y la tabla de hecho, con las jerarquías identificadas en sus respectivas dimensiones.

### Integración de datos

Una vez construido el modelo lógico se procede a poblarlo con datos, utilizando técnicas de limpieza, calidad de datos y procesos ETL. Posteriormente se definieron las reglas y políticas para su respectiva actualización, así como también los procesos que la llevarán a cabo utilizando la herramienta Pentaho Data Integration (PDI).

La estrategia de poblado para todas las dimensiones pasa por capturar el valor de sus atributos de la base de datos del sistema GLPI o generarlos en caso de no existir como, por ejemplo: los campos nombres de las dimensiones `dim_tracking_priority` o `dim_tracking_requesttype`. Después de capturar los datos se modifican los valores nulos o tratan las incongruencias y como paso final se insertan o actualizan en la tabla dimensión en dependencia del tipo SCD —Dimensiones Lentamente Cambiantes— definido para cada campo previamente.

El proceso comienza con la inserción o actualización de los datos de las dimensiones para entonces proceder a poblar la tabla de hecho. Esta tarea finaliza con una transformación encargada de enviar mensajes a los administradores de DW cuando el trabajo fue procesado con éxito.

Además, se definen las siguientes políticas de carga y actualización del almacén de datos tomando en cuenta los tiempos de ejecución de los ETL en base a generar la menor afectación sobre el servicio:

La información se refrescará todos los días a las 12:10 de la noche y se cargarán los datos relativos al día anterior.

Los datos de las dimensiones `dim_affected_service`, `dim_solutiontype`, `dim_entity`, `dim_user`, `dim_tracking_state`, `dim_tracking_category`, `dim_tracking_level`, `dim_tracking_requesttype`, `dim_tracking` y `dim_utilcategories` serán incrementados o actualizados en cada ejecución de la tarea “MASTER”.

Los datos de la dimensión “`dim_date`” serán incrementados en correspondencia a la necesidad de incluir nuevos años.

### Cubos OLAP

En los epígrafes anteriores se definió el modelo lógico del DW y se construyeron los ETL encargados de

poblar las dimensiones y la tabla de hechos, por tanto, solo resta realizar el diseño y publicación de los cubos OLAP en el servidor de Pentaho, para que puedan ser utilizados por las herramientas de análisis.

Con este objetivo se utiliza la aplicación Schema Workbench 5.2 que forma parte de la suite Pentaho.

Para el análisis de la información se pueden definir tantos cubos como sean necesarios, donde cada uno estaría orientado a un tipo de análisis en particular. Por ejemplo, se crearon dos cubos definidos de la misma tabla de hechos pero con propósitos diferentes.

**Cubo Tiques Cerrados:** cubo general que incluye la mayoría de las dimensiones y medidas obtenidas en pasos anteriores.

**Cubo Niveles Atención:** cubo centrado en el análisis de la eficiencia de la mesa de ayuda y el escalonado de los tiques.

### Reportes y Tableros de mando

La herramienta usada para la creación de reportes es Pentaho Report Designer, un editor basado en eclipse destinado a desarrolladores que permite construir y personalizar informes acordes a las necesidades de información.

Para el desarrollo de los tableros de mando se utilizó Community Dashboard Editor (CDE), instalada como parte de CTools: conjunto de herramientas de código abierto desarrollado por Webdetails.

De forma general, se desarrollaron 10 reportes y 4 cuadros como, por ejemplo: eficiencia en las Operaciones. Este cuadro de mando tiene dos vistas una nacional y una territorial. Mediante el mismo se puede comprobar a simple vista cómo se está comportando la Gestión de Tiques en las distintas mesas de ayuda del país.

### Validación de la arquitectura

Con el objetivo de valorar la efectividad de la arquitectura desarrollada se aplicó una encuesta basada en el modelo desarrollado por González (2012) para medir el impacto del Data Warehouse y la Inteligencia de Negocio en las empresas.

En el estudio antes mencionado el autor propone el empleo de 29 variables que responden a 6 constructos: Calidad de la Información, Calidad del Sistema, Calidad del Servicio, Uso, Satisfacción del Usuario e Impacto Individual. Las medidas utilizadas para medir cada uno de los ítems de los constructos son medidas perceptuales.

Se toma en cuenta que los métodos más conocidos para medir por escalas las variables que constituyen actitudes son el método de escalamiento de Likert, el diferencial semántico y la escala de Guttman (Hernández, Fernández, y Baptista, 2014). Para la confección del cuestionario se emplearon 11 afirmaciones, que se evalúan utilizando una escala de Diferencial Semántico de 5 puntos y 8 afirmaciones que utilizan una escala de Likert de 5 puntos asociadas a los constructos Calidad de la Información, Calidad de la Arquitectura, Impacto Individual e Impacto en la Operación de Servicios TI.

La encuesta se sometió a consideración de directivos de la empresa, vinculados al proceso de Tecnologías de la Información y en particular a la operación de los servicios.

En la selección se tuvo en cuenta que estuvieran representados todos los actores y roles que interactúan con la Arquitectura en los diferentes niveles de la actividad de TI. Entre los que se encuentran: directores, jefes de departamentos, jefes de grupos de soporte y especialistas vinculados al desarrollo de aplicaciones BI. En total, se recogieron criterios de 18 sujetos con más de un año de experiencia en el cargo.

Para que realizaran sus valoraciones se les presentó la Arquitectura detallando su funcionamiento, las características, el alcance de los cuadros de mandos y los reportes desarrollados. Se les explicó cómo utilizar los cubos en las vistas de análisis y después de 3 meses de uso se les aplicó el instrumento diseñado. Los datos fueron procesados haciendo uso del software estadístico IBM - SPSS.

Por tanto, después de procesar todas las encuestas se obtuvo que 77,78% de los sujetos encuestados coincidieron que la calidad de la Información es “Muy Buena” y 22,22% restante “Buena”, medida a través 6 afirmaciones: relevancia (utilidad y aplicabilidad) de los datos, nivel de detalle de los datos, exactitud de los datos, actualidad de los datos, comprensión de los datos y los datos que provee el sistema están completos (todo lo que se necesita).

El 50% de los encuestados opina que la calidad de la Arquitectura es “Muy Buena” y el restante 50% “Buena”, se emplearon 5 afirmaciones relativas a la facilidad para ubicar los datos, acceso al sistema, herramientas de acceso a los datos, tiempo de espera y respuesta de la arquitectura y flexibilidad de la arquitectura para cambiar frente a nuevas exigencias.

Al medir el Impacto Individual se obtuvo 61,11% “Buena” y 38,89% “Muy Buena”, algo parecido sucede con el Impacto en la operación de servicios TI con resultados de 72,22% como “Buena” y un 27,78% “Muy Buena”.

De forma general, se puede evaluar la efectividad de la arquitectura Data Warehousing desarrollada entre “Buena” y “Muy Buena” después de procesar todas las encuestas como se muestra en la figura 2.

Sin embargo, un análisis de los descriptivos estadísticos indica la necesidad de mejorar los tiempos de respuesta de la Arquitectura, incrementar los indicadores a analizar, capacitar a los usuarios y mejorar las herramientas para el análisis de datos.

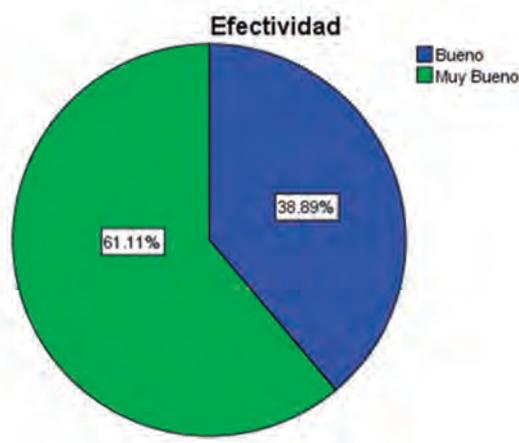


Figura 2. Efectividad Arquitectura DW

### Conclusiones

En esta investigación se ha analizado el papel de la gestión de las Tecnologías de la Información y su importancia en el mundo empresarial. La caracterización de los servicios de TI permitió identificar indicadores y métricas asociadas a la operación de las TI, además de constatar la necesidad de desarrollar una arquitectura como base para soluciones de Inteligencia de Negocio, que velen por la calidad de los servicios.

El estudio de los referentes teóricos relativo a las tecnologías para el desarrollo de arquitecturas Data Warehousing propició la selección del enfoque, metodología y las herramientas a utilizar. Siguiendo los pasos descritos en la metodología Hefesto, se obtuvo el diseño lógico del almacén de datos partiendo de perspectivas e indicadores definidos a partir de las necesidades información y los requerimientos del grupo de trabajo de TI creado para estos fines.

La implementación de la arquitectura se sustentó en soluciones de software libre, PostgreSQL como gestor de base de datos y las herramientas de la Suite Pentaho BI para garantizar los procesos de extracción, transformación y carga del DW, el análisis multidimensional de los datos, el diseño y publicación de reportes y cuadros de mando.

En consecuencia, se desarrolló una arquitectura Data Warehousing que permitió monitorear indicadores de la operación de los servicios de TI, brindando a los directivos la posibilidad de analizar el comportamiento y tendencia de los mismos en el tiempo, y actuar en consecuencia para mejorar la calidad del servicio. Cumpliendo de esta forma el objetivo general de la investigación.

## Referencias

- Aranibar, J. C. (2013). *Sistemas de Información Gerencial para la Administración del Desempeño Empresarial. La Paz, Bolivia: Gráfica Holding srl.*
- Bernabeu, R. D. (2010). *Data Warehousing: Investigación y Sistematización de Conceptos. Libro Electrónico <http://www.dataprix.com/es/data-warehousing-hefesto>, Cordova, Argentina.*
- Berson, A., y Smith, S. J. (1997). *Data warehousing, data mining, and OLAP.* McGraw-Hill, Inc.
- Brizuela, E. I. L., y Blanco, Y. C. (2013). Metodologías para desarrollar Almacén de Datos. *Revista Arquitectura e Ingeniería*, 7(3), 3–12.
- Calderón, C. A., y Pompa, A. S. (2012). El Mapa de Operaciones de las Telecomunicaciones mejorado, eTOM. *Revista Telemática*, 10(2).
- Cocanguilla, S., y Santino, P. (2016). *Diseño e implementación de un modelo para el control de las rutinas de operación y mantenimiento de acuerdo a las mejoras prácticas ETOM, ITIL Y COBIT para la red IP/MPLS de la Corporación Nacional de Telecomunicaciones.* (B.S. thesis). Universidad de las Fuerzas Armadas ESPE. Carrera de Ingeniería Electrónica en Redes y Comunicación de Datos.
- Curto, J. (2010). *Introducción al business intelligence.* Barcelona, España: Editorial UOC.
- Davenport, T. H., y Harris, J. G. (2007). *Competing on analytics: The new science of winning.* Harvard Business Press.
- El-Sappagh, S. H. A., Hendawi, A. M. A., y El Bastawissy, A. H. (2011). A proposed model for data warehouse ETL processes. *Journal of King Saud University-Computer and Information Sciences*, 23(2), 91–104.
- Falcony, J. A. (2013). Modelo de referencia para el mejoramiento de procesos en el área de diseño y calidad de redes de acceso inalámbrico aplicado a empresas de telefonía móvil del suroccidente colombiano integrando las buenas prácticas de eTOM, ITIL y COBIT.
- Garcés, M. B. (2015). *Estudio comparativo de metodologías e implementación de alternativas business intelligence opensource vs. propietarias en entornos tradicionales; caso prototipo en las pymes en el sector agroindustrial.* Quito: Universidad de las Américas, 2015. Recuperado a partir de <http://200.24.220.94/handle/33000/2660>
- González, R. A. (2012). *Impacto de la Data Warehouse e Inteligencia de Negocios en el Desempeño de las Empresas: Investigación Empírica en Perú, como País en Vías de Desarrollo.* Universitat Ramon Llull.
- Hernández, R., Fernández, C., y Baptista, P. (2014). *Metodología de la investigación* (Vol. 6). Mcgraw-hill México.
- Inmon, W. H. (2005). *Building the data warehouse.* John wiley & sons.
- Jörg, T., y Dessloch, S. (2009). Formalizing ETL Jobs for Incremental Loading of Data Warehouses. En *BTW* (pp. 327–346).
- Kimball, R., y Caserta, J. (2004). *The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data.* John Wiley & Sons.

- Kimball, R., Reeves, L., Ross, M., y Thornthwaite, W. (1998). *The data warehouse lifecycle toolkit: expert methods for designing, developing, and deploying data warehouses*. John Wiley & Sons.
- Kimball, R., y Ross, M. (2013). *The data warehouse toolkit: The definitive guide to dimensional modeling*. John Wiley & Sons.
- Kimball, R., Ross, M., y others. (2002). The data warehouse toolkit: the complete guide to dimensional modelling. *Nachdr.J. New York [ua]: Wiley*.
- Kimball, R., Ross, M., Thornthwaite, W., Mundy, J., y Becker, B. (2008). *The Data Warehouse Lifecycle Toolkit*. John Wiley & Sons. Recuperado a partir de <http://dl.acm.org/citation.cfm?id=1571714>
- Mabilon, A. (2010). Retos de la gestión de las tecnologías de información en una nueva era de colaboración social. *Nuevas tendencias en management: Fundamentos y aplicaciones*, 97–125.
- Montalvillo, L. (2012). *Definición y desarrollo de herramienta web de gestión de metadatos Business Intelligence*. Universitat Politècnica de Catalunya.
- Mujica, J. F., y Arzola, M. (2007). Análisis comparativo entre eTOM e ISO 9001: 2000 para empresas de telecomunicaciones. En *XI Congreso de Ingeniería de Organización* (pp. 0185–0194).
- Mujica, J. F., y Arzola, M. (2008). Análisis comparativo de modelos con enfoques basado en procesos para empresas de Telecomunicaciones. En *II International Conference on Industrial Engineering and Industrial Management* (pp. 105–114).
- Paredes, B. R. (2014). *Herramienta de control estratégico de la tecnología*. Universidad de las Fuerzas Armadas ESPE. Maestría en Planificación y Dirección Estratégica.
- Peralta, V. (2001). *Diseño lógico de Data Warehouses a partir de esquemas conceptuales multi-dimensionales*. Master Thesis. Universidad de la República, Uruguay.
- Pérez, M. (2015). *Business Intelligence. Técnicas, herramientas y aplicaciones*. RC Libros.
- Piattini, M. G., y Hervada, F. (2007). *Gobierno de las tecnologías y los sistemas de información*.
- Porter, M. E. (2015). *Ventaja competitiva: creación y sostenimiento de un desempeño superior*. Grupo Editorial Patria.
- Rodríguez, C., y Almarcha, C. (2006). Agilidad empresarial y tendencias en los sistemas de información de los operadores de telecomunicaciones. *Economía Industrial*, (361), 81–91.
- Stratebi, T. (2017). Comparativa de herramientas Business Intelligence. Recuperado 4 de julio de 2017, a partir de <http://todobi.blogspot.com/2017/04/comparativa-de-herramientas-business.html>
- Sunthonwutinun, W., y Chooprayoon, V. (2013). A benchmarking study of standard frameworks for information technology governance. En *The Second Asian Conference on Information Systems, ACIS*.
- Trujillo, J., y Song, I.-Y. (2008). *New trends in data warehousing and OLAP*. North-Holland.
- Turban, E., Volonino, L., y Pollard, C. (2008). *Information technology for management: Transforming organizations in the digital economy*.

