

# Modelo de Datawarehouse para el sector Público (Caso Editora Perú S.A.)

Angel Hermoza Salas

Universidad Nacional Mayor de San Marcos

angelhermozasalas@yahoo.es

## Resumen

*En este paper se presenta y describe un modelo general Olap y prototipo de un Sistema DataWareHouse para una empresa del sector público en general y se implementa en la empresa pública Editora Perú S.A. Se revisan los antecedentes, cómo se consolida la información actualmente de forma manual o con apoyo de otros sistemas, se define el problema, se muestra gráficamente la situación actual, se determina la justificación del presente trabajo y los métodos utilizados. Se detallan los objetivos generales y específicos; además se explica el concepto de Inteligencia de Negocios y Almacén de Datos. Se muestra el Modelo General OLAP de Editora Perú, así como el prototipo desarrollado para mostrar parte de la solución al problema. Finalmente se expone las conclusiones, las recomendaciones y trabajos futuros.*

## 1. Introducción

La generación de reportes detallados, resumidos y comparativos son el medio más utilizado para explotar la información del Sistema ERP Baan, un sistema ERP (Enterprise Resource Planning), según Kwon [Kwon 2001], es “un paquete de software amplio integrado empresarial diseñado para mantener los más altos estándares de calidad de los procesos empresariales”. Para analizar la información los usuarios mensualmente generan un archivo en formato ASCII con la información de los reportes, luego mediante el procedimiento FTP (File Transfer Protocol) se importa el archivo del ambiente AIX al ambiente Windows (computador del usuario), luego desde Microsoft Office se procede a la importación y formateo del archivo, para finalmente preparar los cuadros y gráficos que serán entregados a los ejecutivos de la empresa para la toma de decisiones.

### Definición del problema

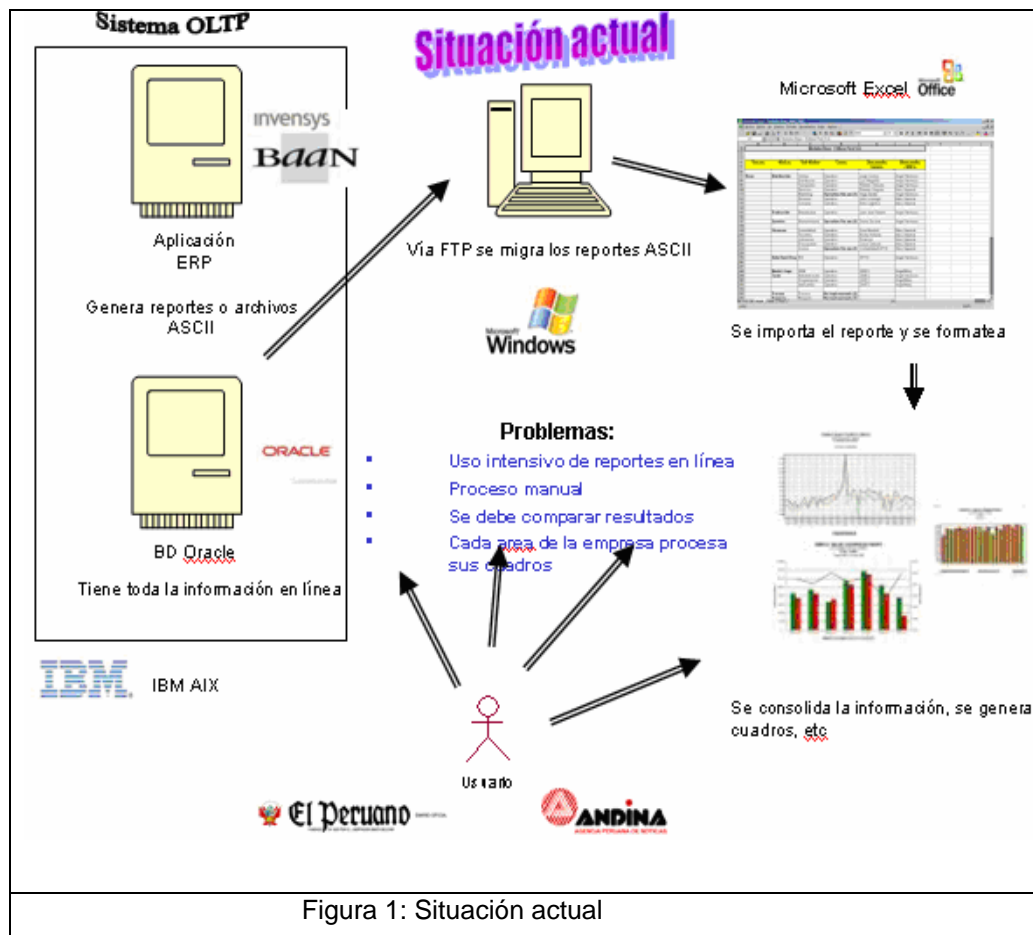
Editora Perú S.A. es una empresa del sector público que ha implementado el sistema ERP Baan IV desde el año 2001, donde todo el procesamiento de transacciones en línea (OLTP) es soportado por este sistema. Como se cuenta con información histórica en el repositorio o base de datos (ORACLE), es necesario procesar, explotar, analizar, para informar a todos los niveles de decisión de la empresa. Actualmente este proceso se realiza con un uso intensivo de procesamiento en línea, el uso de herramientas diversas, el tiempo de consolidación y finalmente los resultados deben ser validados debido a que en el transcurso del proceso pueden haber variado. En la Figura 1 se muestra la situación actual: El Sistema ERP Baan se utiliza para generar reportes históricos que consume recursos en línea ocasionando lentitud y compitiendo con el trabajo diario. Cada fin de mes se solicita la generación de información consolidada, no se atiende a tiempo debido a que el procedimiento para su realización es tedioso y poco confiable.

Los problemas son los siguientes:

- El uso intensivo de reportes en línea para la generación de reportes históricos compite con el procesamiento de transacciones del día, lo que ocasiona lentitud en el tiempo de respuesta del sistema.
- El procesamiento de la información genera archivos ASCII, los que son formateados en Microsoft Office Excel, no garantiza la integridad debido a la manipulación de esta información.
- El ingreso de transacciones que afectan el periodo procesado ocasiona que hayan diferencias

entre lo procesado y la información real.

- Debido al procesamiento manual no se cuenta con la información en la fecha acordada.
- No hay estandarización de los cuadros y gráficos presentados como parte del análisis.



### Limitaciones de la solución del problema

No se cuenta con partida presupuestal para el desarrollo del sistema, por lo que no se puede contratar a terceros, tampoco comprar una solución a medida, ni compra de licencias ni hardware. Solamente se cuenta con el recurso humano del Departamento de Desarrollo para solucionar el problema.

### Variantes de la solución del problema

De contar con presupuesto la solución del problema puede encargarse a un tercero para un desarrollo a medida o mediante la compra de una solución a una empresa de software.

### Justificación e importancia

- La falta de información consolidada no permite tomar decisiones rápidas, solamente se cuenta con información mensual.
- El tiempo de procesamiento de la información ocasiona que no se cumpla con la entrega oportuna.
- No existe un solo repositorio de información dedicado exclusivamente a explotar la información histórica.
- No existe una herramienta de análisis que permita acceder a la información histórica.

### Objetivos Generales

Proponer un Modelo General OLAP para Editora Perú S.A. que sirva como guía para la elaboración paulatina de los datamarts, que permitirán atender a futuro los requerimientos de información de las diferentes áreas de la empresa.

### Objetivos Específicos

Implementar un datamart de Ventas a partir del Modelo OLAP Sunat que permita solucionar el problema actual.

Utilizar la experiencia y conocimiento adquiridos en la solución del problema para la implementación de los demás datamarts definidos en el Modelo General.

El resto de éste paper está organizado de la siguiente manera. En la sección 2 se muestra como realizar la sección de Trabajos Previos. La sección 3 describe el Modelo Olap General. La forma de colocar los Experimentos y Resultados se encuentra en la sección 4. La Discusión de los Experimentos se muestra en la sección 5 y finalmente, la manera de redactar las conclusiones y recomendaciones o trabajos futuros está en la sección 6.

## **2. Trabajos Previos**

Business Intelligence (Inteligencia de Negocios)

Según [Microsoft, 2004] vivimos en una época en que la información es la clave para obtener una ventaja competitiva en el mundo de los negocios. Para mantenerse competitiva una empresa, los gerentes y tomadores de decisiones requieren de un acceso rápido y fácil a información útil y valiosa para la empresa. Una forma de solucionar este problema es por medio del uso de Business Intelligence o Inteligencia de Negocios. La Inteligencia de Negocios o Business Intelligence (BI) se puede definir como el proceso de analizar los bienes o datos acumulados en la empresa y extraer una cierta inteligencia o conocimiento de ellos. Dentro de la categoría de bienes se incluyen las bases de datos de clientes, información de la cadena de suministro, ventas personales y cualquier actividad de marketing o fuente de información relevante para la empresa. La clave para BI es la información y uno de sus mayores beneficios es la posibilidad de utilizarla en la toma de decisiones. Tal vez le ayude a comprender mejor el concepto por medio de un ejemplo. Una franquicia de hoteles a nivel nacional que utiliza aplicaciones de BI para llevar un registro estadístico del porcentaje promedio de ocupación del hotel, así como los días promedio de estancia de cada huésped, considerando las diferencias entre temporadas.

Con esta información se puede:

- Calcular la rentabilidad de cada hotel en cada temporada del año
- Determinar quién es su segmento de mercado
- Calcular la participación de mercado de la franquicia y de cada hotel
- Identificar oportunidades y amenazas

Estas son sólo algunas de las formas en que una empresa u organización se puede beneficiar por la implementación de software de BI, hay una gran variedad de aplicaciones o software que brindan a la empresa la habilidad de analizar de una forma rápida por qué pasan las cosas y enfocarse a patrones y amenazas.

¿Qué se puede hacer con Business Intelligence (BI)?

Con Business Intelligence (BI) se puede:

- Generar reportes globales o por secciones
- Crear una base de datos de clientes
- Crear escenarios con respecto a una decisión
- Hacer pronósticos de ventas y devoluciones
- Compartir información entre departamentos
- Análisis multidimensionales
- Generar y procesar datos

- Cambiar la estructura de toma de decisiones
- Mejorar el servicio al cliente

La siguiente es una lista de las áreas más comunes en las que las soluciones de inteligencia de negocios son utilizadas:

- Ventas: Análisis de ventas; Detección de clientes importantes; Análisis de productos, líneas, mercados; Pronósticos y proyecciones.
- Marketing: Segmentación y análisis de clientes; Seguimiento a nuevos productos.
- Finanzas: Análisis de gastos; Rotación de cartera; Razones financieras.
- Manufactura: Productividad en líneas; Análisis de desperdicios; Análisis de calidad; Rotación de inventarios y partes críticas.
- Embarques: Seguimiento de embarques; Motivos por los cuales se pierden pedidos.

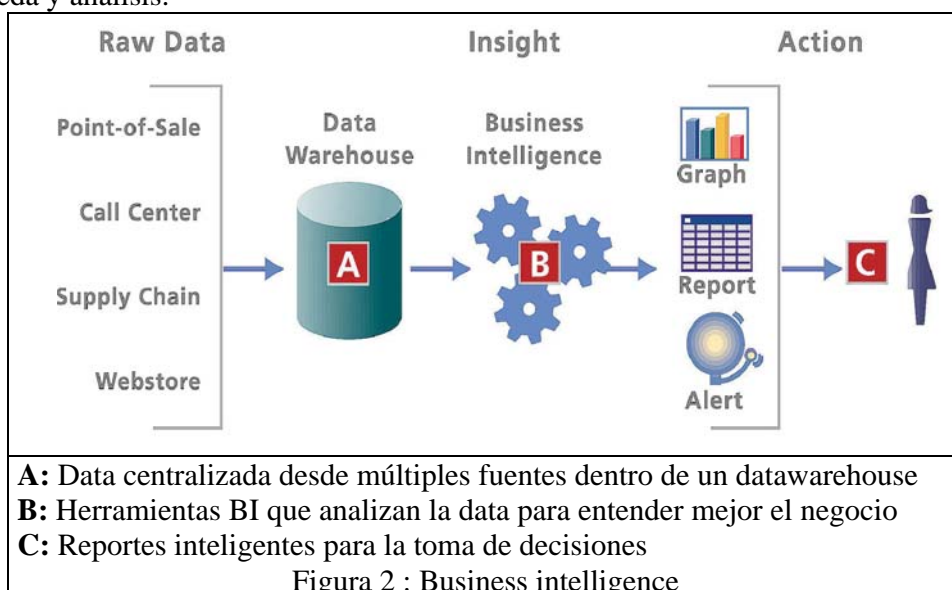
#### Componentes de Business Intelligence

**Multidimensionalidad:** la información multidimensional se puede encontrar en hojas de cálculo, bases de datos, etc. Una herramienta de BI debe ser capaz de reunir información dispersa en toda la empresa e incluso en diferentes fuentes para así proporcionar a los departamentos la accesibilidad, poder y flexibilidad que necesitan para analizar la información. Por ejemplo, un pronóstico de ventas de un nuevo producto en varias regiones no está completo si no se toma en cuenta también el comportamiento histórico de las ventas de cada región y la forma en que la introducción de nuevos productos se ha desarrollado en cada región en cuestión.

**Data Mining (Minería de Datos):** Las empresas suelen generar grandes cantidades de información sobre sus procesos productivos, desempeño operacional, mercados y clientes. Pero el éxito de los negocios depende por lo general de la habilidad para ver nuevas tendencias o cambios en las tendencias. Las aplicaciones de data mining pueden identificar tendencias y comportamientos, no sólo para extraer información, sino también para descubrir las relaciones en bases de datos que pueden identificar comportamientos que no son muy evidentes.

**Agentes:** Los agentes son programas que "piensan". Ellos pueden realizar tareas sin necesidad de intervención humana. Por ejemplo, un agente pueden realizar tareas complejas, como elaborar documentos, establecer diagramas de flujo, etc.

**Data Warehouse (Almacén de Datos):** Es la respuesta de la tecnología de información a la descentralización en la toma de decisiones. Coloca información de todas las áreas funcionales de la organización en manos de quien toma las decisiones. También proporciona herramientas para búsqueda y análisis.



## El Futuro

La Figura 2, describe la Inteligencia de Negocios (BI) que ya no puede ser ignorada por ninguna organización que reconoce que estamos en la era de la información. En el futuro cercano debemos esperar lo siguiente:

- Proyectos más frecuentes y más largos
- Barreras de entrada para los primeros
- La infraestructura será estándar
- Se establecerán centros de información
- Convergencia de tecnologías (acceso por internet)
- Cambiar actividades consideradas periféricas en Core Business

## Soluciones Datawarehouse

Según [Cognos, 2002] en AFP Nueva Vida las soluciones cognos cambiaron la manera como se hacían los negocios. se permitió efectuar seguimiento más a detalle del comportamiento de los afiliados y las empresas aportadoras. asimismo, se identificaron segmentos importantes de clientes y empresas.

Según [Nakasone, 2004] las empresas han optado por utilizar la inteligencia de negocios y el primer escalón es construyendo un almacén de datos (datawarehouse), para luego avanzar en la minería de datos (datamining).

### **3. Modelo Olap para una empresa pública**

A fin de alcanzar los objetivos planteados en el presente trabajo se ha planteado el modelamiento de una solución general para toda empresa del estado, tomamos como base la empresa pública Editora Perú, pero puede extenderse la solución a cualquier otra empresa tal como Sedapal, Essalud, Ministerios, etc.

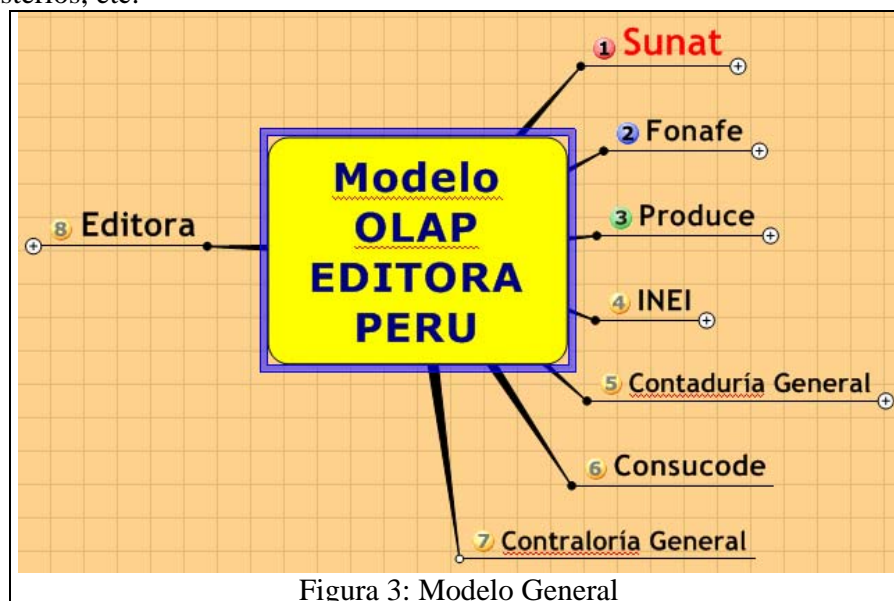


Figura 3: Modelo General

En la Figura 3 se muestra el Modelo General en este caso mostrando a Editora Perú, pero se puede generalizar para toda empresa del estado, todas las empresas tienen la obligación legal de reportar a las siguientes entidades :

- Sunat (Superintendencia Nacional de Administración Tributaria)
- Fonafe (Fondo Nacional de Financiamiento de la Actividad Empresarial del Estado)
- Produce (Ministerio de la Producción)

- INEI (Instituto Nacional de Estadística e Informática)
- Contaduría General de la República
- Contraloría General de la República
- Consucode (Consejo Superior de Contrataciones y Adquisiciones del Estado)

Además internamente Editora Perú debe generar información interna para sus órganos de supervisión y control : Recursos Humanos, Costos, Presupuesto, Indicadores, Gestión Contable, Clientes, Proveedores, Tesorería, Activo Fijo, Manufactura, Distribución.

En la Figura 4 se muestra el Modelo General Detalle, donde se puede observar que cada entidad exige información detallada para fines de control, las que se deben entregar en tiempo y plazos establecidos legalmente.

En la Figura 4 en el extremo superior derecho se muestra el Modelo OLAP Sunat, se resalta el Registro de Ventas que será lo que se desarrollará como parte de este trabajo, en este caso se tiene la base legal, los formularios, formatos de archivo a informar, archivo de transferencia, luego a través del Sistema PDT la información será enviada a la Sunat.

El presente trabajo permite realizar un “reciclaje” o “reuso” de ésta información para alimentar los datamart y conformar el datawarehouse de una empresa pública.

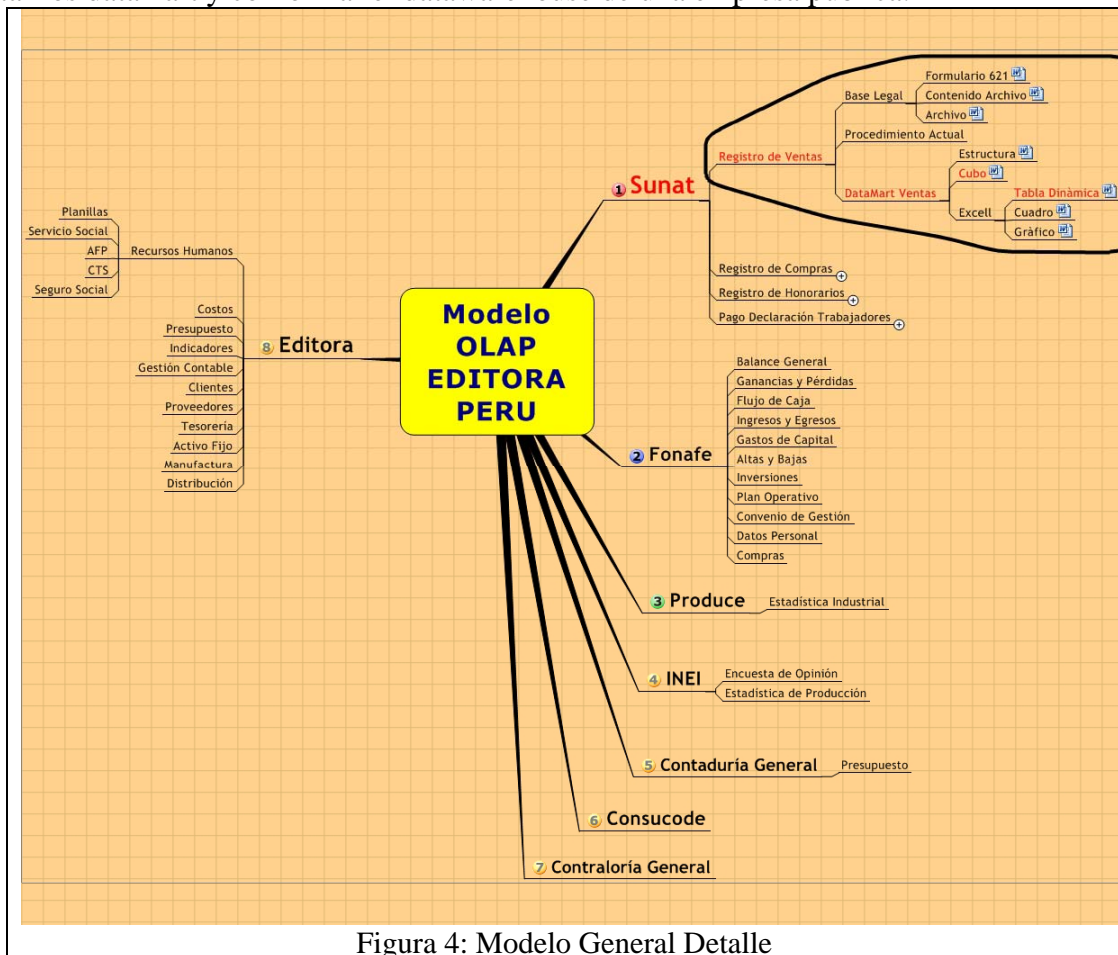


Figura 4: Modelo General Detalle

#### 4. Experimentos y Resultados

##### Organización e instancias de prueba

Se ha considerado los años del 2002 al 2005. Los componentes de hardware con los que trabajará el sistema son: Servidor compatible, Servidor IBM X Series Modelo H70, SAN Storage HP.

Mientras que los componentes de software con los que trabajará el sistema son: Sistema Operativo: Windows 2000 para el servidor, Explorador de Internet (Internet Explorer), Manejador de base de datos MSSQL 2000, Sistema ERP, Sistema Operativo AIX v 4.3, Manejador de Base de Datos Oracle v.9, Sistema ERP BAAN IV C2.

Procesamiento

Se ha desarrollado procedimientos almacenados que permiten transferir los datos del datamart de ventas de la base de datos Oracle v9.0 a Ms Sql 2000, luego otro procedimiento almacenado genera el cubo o datamart de ventas dentro del Analysis Server, en ese momento el usuario tiene disponible la información que se muestra.

Resultados

El usuario debe acceder a la hoja de cálculo y mediante un procedimiento seleccionar el cubo de ventas.

La Figura 5 nos muestra las ventas totales de los años 2002 al 2005 donde podemos visualizar cada mes como se comportan las ventas en Editora Perú.

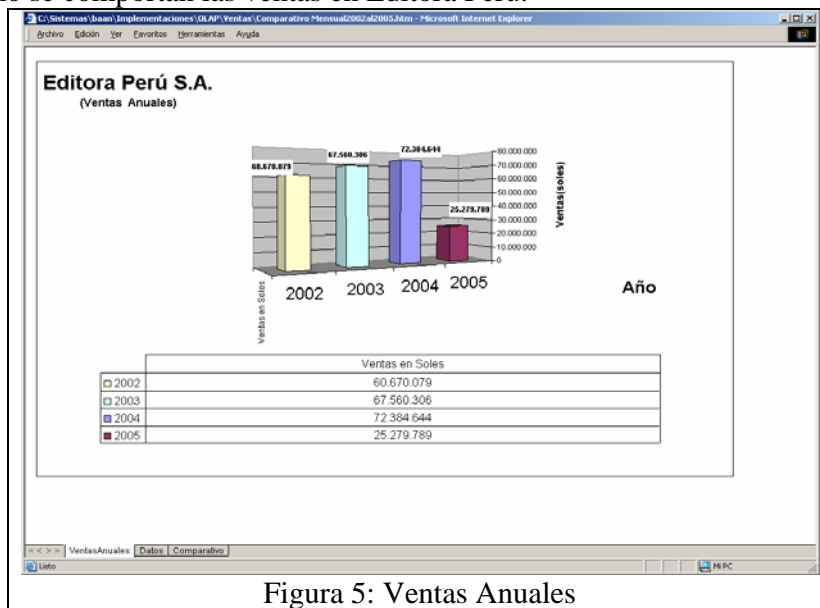


Figura 5: Ventas Anuales

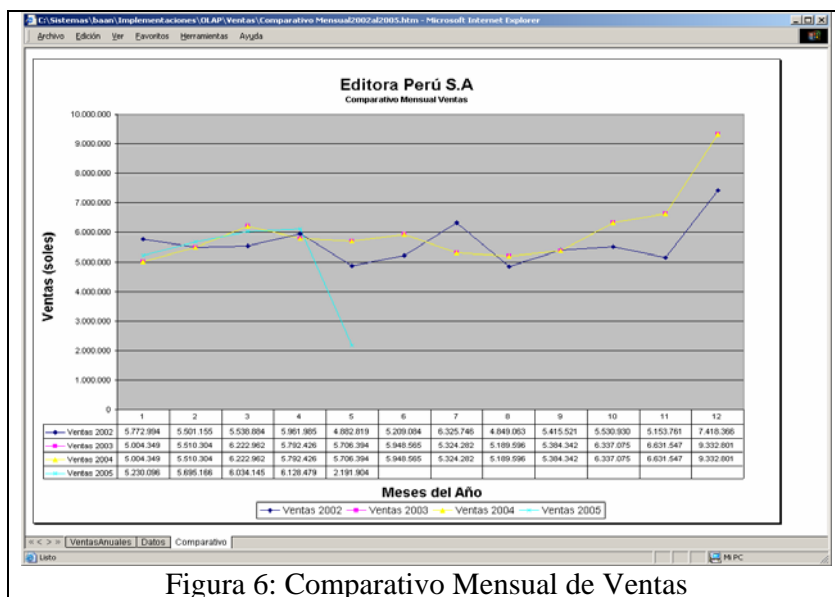


Figura 6: Comparativo Mensual de Ventas

La Figura 6 nos muestra el cuadro comparativo mensual de ventas, en él se puede comparar la venta mensual como anual, de modo que se muestren las tendencias y las expectativas de venta de acuerdo al pronóstico realizado al inicio del año. Esta información está disponible para toda la empresa.

## 5. Discusión de los Experimentos

En la Figura 7 se muestra gráficamente la solución al problema, así como las ventajas que se obtienen con la implementación de esta solución:

Ahora se cuenta con un servidor OLAP diferente al servidor OLTP por lo que ya no se compite con los recursos diarios, el acceso es más rápido y directamente de la hoja de cálculo Microsoft Office.

El procesamiento es totalmente automatizado, se procesa el periodo deseado, en este caso se ha considerado hacerlo quincenal y mensualmente.

Al existir un solo repositorio OLAP ahora todos los usuarios de la empresa pueden acceder simultáneamente a la información sin problemas de lentitud manteniéndose la integridad.

Se ha estandarizado el uso de plantillas para la presentación de la información tanto de los cuadros como de los gráficos.

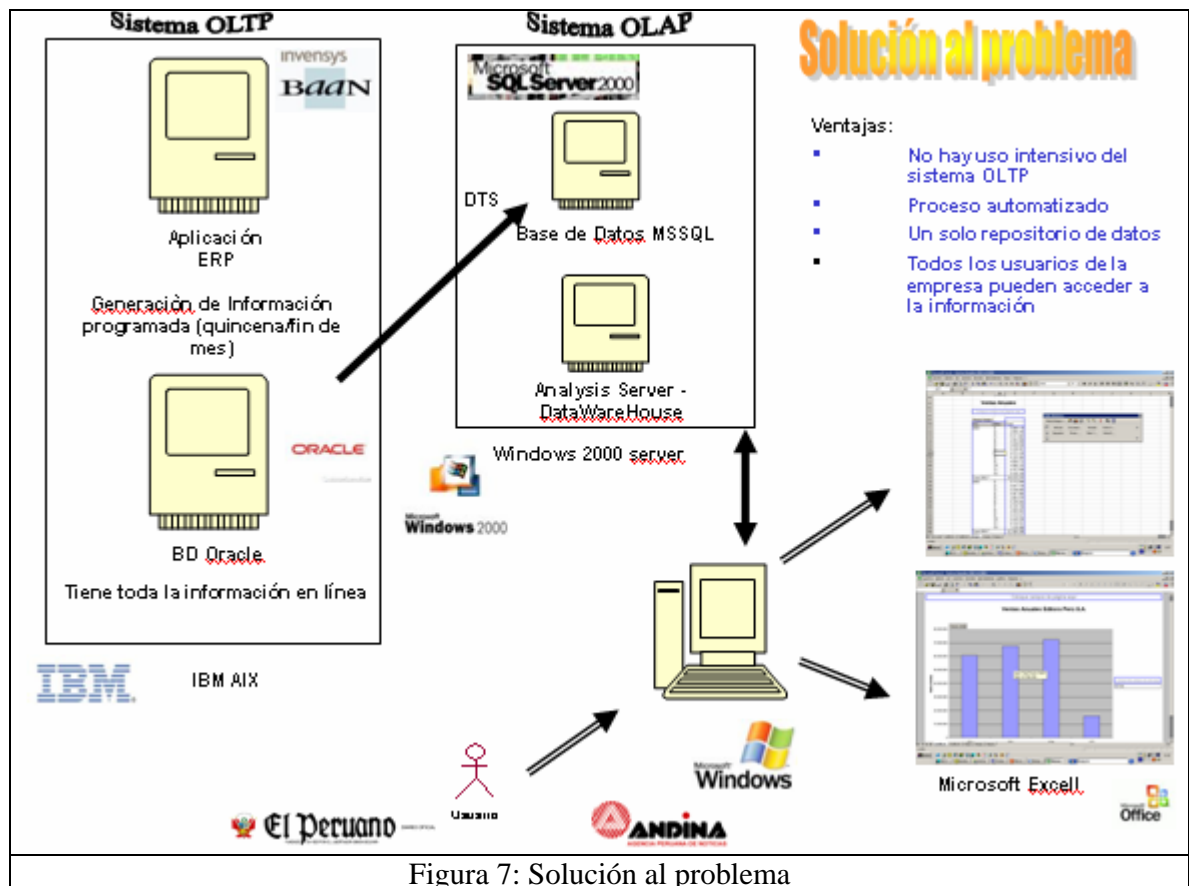


Figura 7: Solución al problema

## 6. Conclusiones

La implementación del datamart de ventas permitirá que el usuario pueda contar con una herramienta en línea totalmente automatizada, de fácil uso, que le permita disminuir el tiempo de procesamiento y dedique mayor tiempo a la etapa de análisis de la información.

Al automatizar la generación de reportes y gráficos:

- Se eliminan los errores o diferencias por migración de datos y formateo.
- Se disminuye el tiempo de procesamiento por procedimientos manuales.

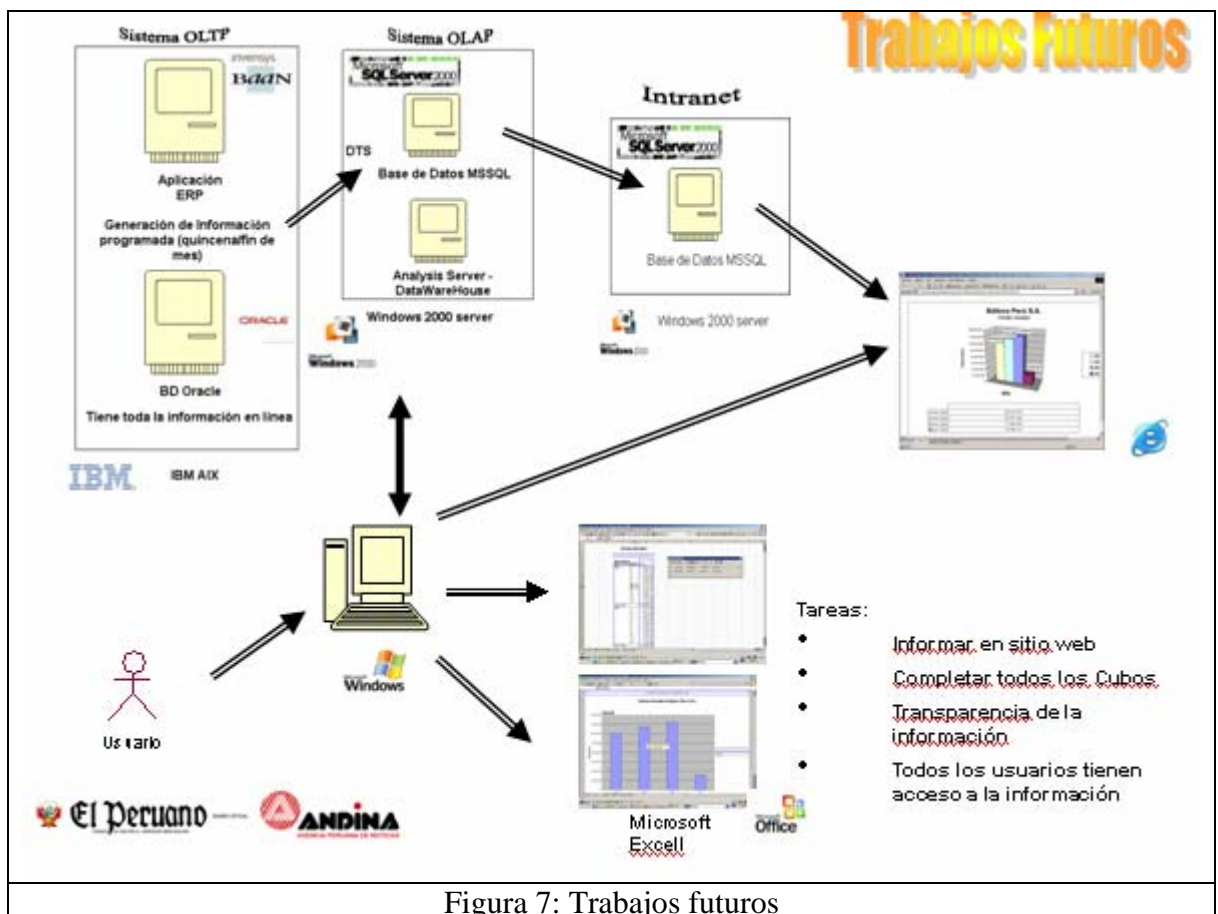
- Por decisión del usuario el procesamiento se realiza quincenal y mensualmente, quedando abierta la posibilidad de efectuarlo diariamente.
- El acceso a la información residente ahora en el servidor OLAP es más rápido en comparación con el acceso a la base de datos del servidor OLTP utilizado en el proceso manual.

Según [Microsoft, 2004], los usuarios pueden acceder a un solo repositorio de datos (servidor OLAP) directamente desde la hoja de cálculo Microsoft Office, utilizando plantillas estándar. El único costo para la implementación de este sistema consiste en la capacitación de la herramienta Analysis Server de Microsoft producto que viene como parte de la Base de Datos MS SQL .

#### Recomendaciones o trabajos futuros

Como se muestra en la Figura 7 existen las siguientes tareas a futuro:

Generar información de los datamarts para informar en el Portal de Editora Perú ([www.editoraperu.com.pe](http://www.editoraperu.com.pe)). Implementar todos los cubos faltantes que se muestran en la Figura 4 Modelo General OLAP Detallado. Que todos los usuarios de la empresa tengan acceso a la información de la empresa vía Intranet.



## Referencias

[Microsoft, 2004] Microsoft (2004). Guía de Estrategia de Business Intelligence. 24 pags, © 2004 Microsoft Corporation.

[Kwon 01] Kwon, O. B. and Leeb, J. J. A multi-agent intelligent system for efficient ERP maintenance, Instituto Handong de Informacion Tecnologica, Expert System and Application, Volume 21, Issue 4, November 2001, Pages 191-202.

[Cognos, 2002] Cognos, Casos de éxito; AFP Unión Vida - Perú, 2003

[Nakasone, 2004] Nakasone Nicolas; OLAP Essential - Línea de Talleres Inteligencia de Negocios, MugPerú, Lima 2004

[Nakasone, 2004] Nakasone Nicolas; OLAP Advanced - Línea de Talleres Inteligencia de Negocios, MugPerú, Lima 2004

[Microsoft, 2004] Microsoft; Designing and Implementing OLAP Solutions with Microsoft SQL Server; 2004, Workbook, Microsoft Training and Certification Microsoft Official  
Microsoft 2004 Crear inteligencia empresarial con Analysis Services de Office XP y SQL 2000

### Otras referencias importantes

[Diane Larsen, 2003] Servicios de transformación de datos (DTS) en Microsoft SQL Server 2000, Microsoft Corporation Site :  
[http://www.microsoft.com/spanish/msdn/articulos/archivo/010201/voices/dts\\_overview.asp](http://www.microsoft.com/spanish/msdn/articulos/archivo/010201/voices/dts_overview.asp), accesado en Enero 2007.

[Cognos, 2005] Delivering Warehouse Return On Investment with Business Intelligence From Cognos®, <http://www.dmreview.com/whitepaper/dwi.pdf>, Datawarehouse Review  
Whitepaper from Sybase, Inc

[Cognos, 2005] Data Warehousing for Healthcare: The Greatest Weapon in Your Competitive Arsenal, <http://www.dmreview.com/whitepaper/WID242.pdf> , Datawarehouse Review  
OLAP Train and Reed Jacobson