



**UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO**

**FACULTAD DE ESTUDIOS SUPERIORES
ACATLÁN**

**Sistema de Análisis de los Registros de Tarificación de
Centrales Telefónicas**

MEMORIA DE DESEMPEÑO PROFESIONAL

QUE PARA OBTENER EL TÍTULO DE
**LICENCIADO EN MATEMÁTICAS APLICADAS
Y COMPUTACIÓN**

PRESENTA

GABRIELA AYALA PÉREZ

ASESOR: ING. ELVIRA BEATRIZ CLAVEL DIAZ

ABRIL DEL 2005

m343861



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

DEDICATORIA

A mi hija,

Quien representa la mayor bendición de mi vida y a quien espero poder guiar por el camino al éxito y a la felicidad.

Autorizo a la Dirección General de Bibliotecas de la UNAM a difundir en formato electrónico e impreso el contenido de mi trabajo recepcional.

NOMBRE: Gabriela Ayala
Perez

FECHA: 6- Mayo - 2005

FIRMA: P.A. Mariana G. Tapalá

DEDICATORIA ESPECIAL

A mi padre,

Por todo el amor, por todo el apoyo, por los buenos momentos, por hacernos reír, porque estoy segura de que, aunque Dios te llamó recientemente a su presencia, sigues cerca de nosotros como el padre maravilloso que conocimos.

AGRADECIMIENTOS

A Dios,

Por la vida, la salud y las bendiciones de cada día.

A mis padres,

Mil gracias por todo el apoyo, porque son parte fundamental de este logro y de todas mis metas. Gracias a mi madre por hacerme saber que tenía que estudiar para salir adelante.

A mis hermanos,

Gracias por todos los años de convivencia, por su compañía y por saber que puedo contar con ustedes al igual que ustedes conmigo. Especialmente a mi hermano Antonio por el ejemplo de valor para afrontar la adversidad.

A mi familia,

Por todos los años de estar a mi lado en momentos difíciles, por el gran cariño y apoyo recibido, particularmente a Jorge, por compartir el amor de nuestra hija y por la gran ayuda que me brindó durante la carrera.

A mis compañeros,

Por la amistad, por todas las cosas que hemos aprendido juntos tanto en la escuela como en el trabajo. En especial mi más sincero agradecimiento a quien creyó en mí y me hizo saber que yo podía lograrlo, que no era un imposible.

A la universidad y a mis maestros,

Por la oportunidad de aprender y tener una profesión que nos abre el camino de la superación.

Sistema de Análisis de los Registros de Tarificación de Centrales Telefónicas

Objetivo General

Automatizar el proceso de análisis de los registros de tarificación de las Centrales Telefónicas, obteniendo los datos de manera remota y proporcionando una interfaz gráfica y de fácil manejo para el usuario, que permita la emisión de reportes claros y confiables en menor tiempo.

ÍNDICE

Objetivo General	5
Introducción	8
Capítulo I. Marco contextual	11
I.1 Descripción del entorno	11
I.1.1 Central Telefónica	12
I.1.2 Características de la Central Telefónica S-12	16
I.1.3 Ubicación de datos	16
I.2 Tarificación	17
I.3 Planteamiento del problema	18
I.4 Las centrales en la WAN / RCD	19
Capítulo II. Análisis de Información	22
2.1 Obtención remota de las bases de datos	23
2.1.1 Nomenclatura de los archivos	23
2.1.2 Recuperación de los archivos	24
2.2 Formato y estructura de los archivos	29
2.2.1 Estructura de los archivos CDF	30
2.2.2 Formato de los archivos CDF	31
2.3 Los registros CDR	31
2.3.1 Estructura de los registros CDR	32
2.3.2 Interpretación de los registros CDR	35
2.4 Información de salida	38
Capítulo III. Propuesta de Solución	42
3.1 Diagnóstico de las circunstancias	42

3.1.1 Identificación de recursos	42
3.1.2 Solución propuesta.....	43
3.1.3 Soluciones alternativas.....	44
3.1.4 Justificación	45
3.2 Descripción de procedimientos.	47
3.2.1 Obtención de los archivos de tarificación	47
3.2.2 Procesamiento de archivos	49
3.2.2.1 Desglose de llamadas.....	50
3.2.2.2 Envío de datos a Excel	53
3.2.2.3 Verificación de secuencia	55
3.3 La interfase gráfica	57
3.4 Reportes resultantes	60
3.4.1 Reporte de Desglose de Llamada	60
3.4.2 Reporte de Verificación de Secuencia	61
3.5 Implementación	63
3.5.1 Proceso de Prueba	63
3.5.2 Instalación de la interfase gráfica en las PCs de los usuarios.....	64
3.5.3 Capacitación de usuarios	66
Capítulo IV. Discusión de Resultados.....	68
4.1 Interpretación de reportes obtenidos.....	68
4.1.1 Ejemplo 1: Central Mayorazgo	69
4.1.2 Ejemplo 2 Central Ocotlán	75
4.2 Mantenimiento del sistema.....	80
Conclusiones.....	82
Anexos	85
Anexo A. Glosario	85
Anexo B. Lista de Figuras.....	86
Bibliografía.....	87

Introducción

Uno de los principales problemas que enfrentan las grandes compañías de telecomunicaciones al generar y asegurar sus ganancias es la *tarificación*, es decir, la capacidad de realizar el cobro por servicio a sus subscriptores de manera efectiva y adecuada. La trascendencia de que las compañías telefónicas cuenten con procesos que aseguren un correcto cobro del servicio que prestan, tiene como consecuencia una mayor confianza por parte de sus clientes. Es por este motivo que, día con día las compañías telefónicas buscan implementar procesos que mejoren la manipulación de su información y que garanticen su fiabilidad.

Por otro lado, el fácil acceso a la telefonía en nuestros días hace que el número de personas que cuentan con este servicio sea cada vez mayor, y representa para las compañías telefónicas grandes volúmenes de información que debe ser procesada en el menor tiempo posible.

En esta memoria se presenta la solución aplicada a un problema de procesamiento de datos que se suscitó en una compañía que pertenece al giro de telecomunicaciones y que opera exclusivamente en el área de "telefonía fija". La solución propuesta involucra el análisis y diseño de un sistema que proporcione datos verídicos y confiables a través de un proceso automático que optimice el tiempo de generación de reportes, aplicando los conocimientos adquiridos en la Licenciatura en Matemáticas Aplicadas y Computación.

En el Capítulo I se describe el entorno en el cual surge la necesidad de automatizar un proceso, el tipo de información que se maneja, la forma en que ésta es almacenada y la utilidad que representa para la empresa.

El objetivo del capítulo II es describir el proceso de obtención de lo que serán nuestros *datos de entrada*, su interpretación así como también la *información de salida* que se espera obtener como resultado.

En el capítulo III se presenta la solución aplicada, la cual consiste en el desarrollo de un sistema que involucra el diseño de una interfase gráfica de usuario y el procesamiento de la información, se muestran también los diagramas de flujo correspondientes a funciones específicas utilizadas en la aplicación, asimismo se describe el proceso de implementación del sistema.

Finalmente, en el capítulo IV se analizan los alcances y limitaciones del sistema después de su implantación y se discuten las características de los reportes resultantes.

Capítulo I. Marco Contextual

Cuando quieres algo y estás realmente convencido, el universo entero conspira para que lo logres

Anónimo

Capítulo I. Marco contextual

Inicialmente debe considerarse que es de suma importancia ubicarnos en el entorno sobre el cual surge la necesidad de automatizar un proceso, describir el tipo de información que se maneja, la forma en que ésta es almacenada y la utilidad que representa para la empresa.

I.1 Descripción del entorno

El proyecto fue desarrollado para Alcatel, que puede describirse como una compañía perteneciente al giro de telecomunicaciones que, en más de 130 países, diseña, desarrolla y construye competitivas redes de comunicaciones, proporciona la tecnología, infraestructura, servicios de instalación y soporte técnico a otras compañías que proveen servicios de transferencia de voz, datos y multimedia para cualquier tipo de cliente en cualquier parte del mundo.

La compañía cuenta con un CNS (Centro Nacional de Servicio) el cual proporciona servicios de operación y mantenimiento a las centrales telefónicas de las redes de telefonía pública. Estas redes operan actualmente dos tipos de telefonía:

- *Telefonía fija*
- *Telefonía móvil*

Por **telefonía fija**, nos referimos a la red de teléfonos instalados en domicilios particulares, comercios y compañías establecidas los cuales están conectados directamente a través de un par trenzado de alambres de cobre a una central telefónica que típicamente está ubicada a no más de 7 Km. a la redonda; por **telefonía móvil** nos referimos a la red de teléfonos también conocidos como "celulares" los cuales no se encuentran conectados físicamente a través de cables a una central pero que se comunican con ellas por medio de señales de radiofrecuencia. Si bien es cierto que es posible la comunicación desde las redes de telefonía fija hacia las redes de telefonía móvil y viceversa, la administración y operación de cada una es independiente.

En esta memoria nos enfocamos exclusivamente a los datos generados por las llamadas realizadas ya sean de voz o datos, utilizando la red de telefonía fija.

I.1.1 Central Telefónica

Para poder establecer una llamada telefónica de un abonado a otro, necesariamente el aparato telefónico debe estar conectado a una Central Telefónica¹. A lo largo del país y para cada localidad, existen centrales que cuentan con la tecnología e infraestructura necesarias para concentrar un conjunto de suscriptores conectados a equipos que permiten establecer la comunicación o transferencia de voz o datos desde un lugar a otro.

¹ Nos referiremos en lo sucesivo a la Central Telefónica como "la central". El término utilizado mundialmente para referir una central es **TELEPHONIC EXCHANGE**.



Figura 1. Red Telefónica Pública Conmutada

La Red Telefónica Pública Conmutada (RTPC) (Ver Figura 1), también conocida por sus siglas en inglés PSTN (Public Switched Telephone Network), es la encargada de suministrar servicios de transmisión de voz a nivel internacional y está formada por un número finito de centrales ubicadas en diferentes localidades e interconectadas entre sí para formar una red, en la que cada central representa un nodo.

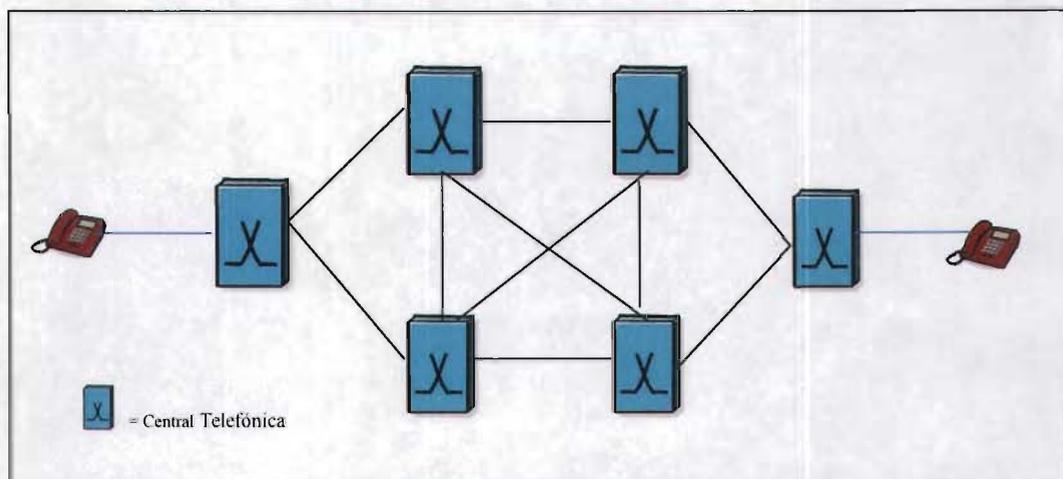


Figura 2. Interconexión de Centrales

Como podemos apreciar en la Figura 2, al realizar una llamada desde un teléfono a otro, es necesario pasar por un cierto número de centrales dependiendo del destino de la llamada. A lo largo de cada país, éstas se encuentran ubicadas de manera que la mayor parte de los habitantes de cada localidad, municipio, colonia, etc. pueda tener acceso a una línea telefónica (ver Figura 3), cabe señalar que la solución aquí presentada aplica exclusivamente para el tipo de centrales telefónicas que fabrica y comercializa Alcatel: S-12² y que operan en México.



Figura 3. La centrales a lo largo del país

Debido a que existen varias compañías de telecomunicaciones que fabrican sus propias centrales, existen también diferentes tipos de centrales que forman una gran variedad de redes que difieren en su funcionamiento y en la manera en que manejan su información. Sin embargo, de acuerdo a las normas y estrategias internacionales, estas redes deben ser compatibles entre sí de modo que cada central tenga la capacidad de comunicarse con

² En nuestro país operan principalmente tres tipos de centrales: AXE, S12 y 5E. Las centrales S12 manejan más del 47% de la telefonía fija, las centrales AXE operan una cantidad similar, el resto lo operan centrales 5E y otras compañías.

las centrales de otras compañías para formar parte de la Red Telefónica Pública Conmutada y garantizar así la comunicación entre todos los nodos de la red.

Cada una de las centrales S-12 está compuesta por diferentes módulos que se encargan de controlar diferentes funciones y así garantizar el establecimiento de la llamada (ver Figura 4). La central cuenta con interfases telefónicas que permiten la comunicación entre *abonados*³ dentro de la red conectándose entre ellos para lograr el intercambio de información (voz y datos). El aspecto de la central es similar al que se muestra en la Figura 4.

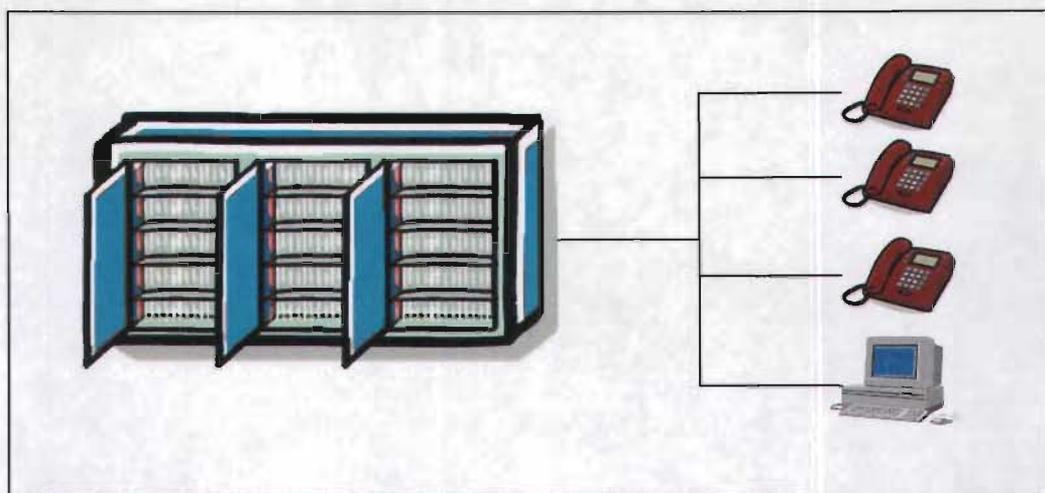


Figura 4. Central Telefónica

³ En términos de telefonía, un abonado es el individuo que se encuentra suscrito a una compañía telefónica y tiene acceso al uso de los servicios que proporciona la RTPC, también puede ser llamado "subscriber".

I.1.2 Características de la Central Telefónica S-12

Las centrales S-12 se caracterizan por dos propiedades esenciales: *tecnología digital* y *control distribuido*. Primero, el S-12 usa tecnología digital porque su control y funciones son realizados por programas que son ejecutados en microprocesadores, y el manejo interno de la información (conmutación y transmisión) se lleva a cabo mediante técnicas totalmente digitales. Segundo, el control distribuido significa que las funciones llevadas a cabo por el sistema están divididas en paquetes de tareas agrupadas en forma homogénea y se asignan a elementos de control específicos.

El software es organizado bajo la plataforma de un sistema operativo propio y una base de datos específica del sistema, con base a éstos se ejecuta un conjunto de programas de aplicación o módulos software. El sistema operativo está compuesto por una serie de funciones software que hacen posible la administración de los recursos del sistema (distribución de tiempos del CPU, administración de la memoria, comunicación a través de la red, etc.). Aunque existen variantes de centrales S-12 con base al tipo de operaciones que realiza cada una, la información de tarificación se almacena indistintamente en cada una de ellas.

I.1.3 Ubicación de datos

Toda la información se guarda en las unidades de almacenamiento o unidades de disco de la misma central (ver Figura 5). En estos discos se encuentra almacenada la base de datos que contiene la información organizada en forma de archivos conteniendo tablas.

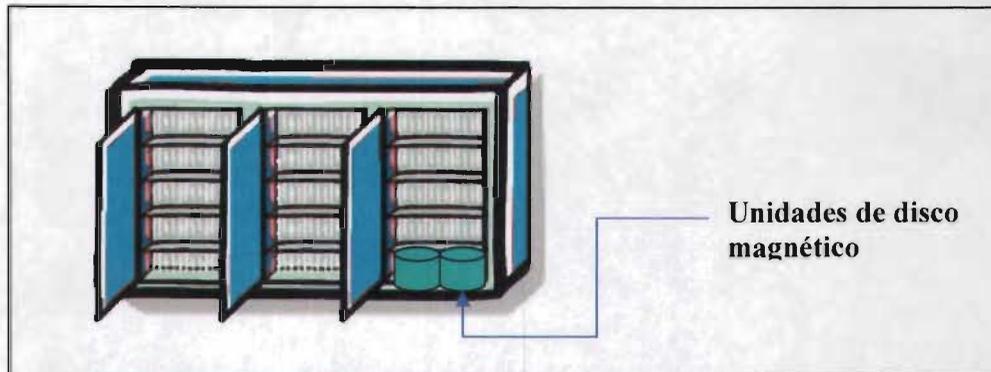


Figura 5. Unidades de disco

De todas las tablas contenidas en la base de datos, las que nos interesan son las que tienen la información correspondiente a tarificación. Cada llamada es almacenada en forma de registro de manera que puedan obtenerse datos importantes sobre la llamada y así iniciar el proceso de tarificación.

I.2 Tarificación

Se define «tarificación» como el proceso de cobro por servicio al suscriptor al utilizar la infraestructura diseñada para el establecimiento de una llamada telefónica. Este proceso complejo se lleva a cabo en la base de datos de la central telefónica, en donde se generan y almacenan todos los datos necesarios para llevar a cabo el análisis y desglose de cada llamada. En primera instancia, la tarificación involucra la colección, adaptación y clasificación de datos, así como la emisión de un recibo para cada suscriptor.

Las funcionalidades del proceso de tarificación dentro de las centrales telefónicas proveen diferentes métodos para tener la capacidad de realizar cargos a los subscriptores al hacer uso de los servicios proporcionados por la red telefónica pública. Cada llamada realizada generará un registro en la base de datos.

En este esquema el proceso de manipulación de los datos resulta lento y tedioso debido a la gran cantidad de información que se registra diariamente en la base de datos de la central.

I.3 Planteamiento del problema

En el CNS se reciben y atienden todas las quejas y solicitudes por parte del cliente con relación al funcionamiento de las centrales, continuamente se reciben quejas acerca de faltantes de registros de llamadas en determinadas centrales en un rango de tiempo específico, para poder dar una respuesta apropiada al cliente es necesario recuperar y analizar los archivos de tarificación generados por la central.

El proceso de manipulación de registros de tarificación, se realizaba manualmente usando un editor de textos de gran capacidad (SPF) para procesarlos y posteriormente trasladarlos a una hoja de cálculo "Excel", lo cual toma un tiempo considerable debido al tamaño y formato de los archivos, tomando en cuenta que en muchas ocasiones es necesario procesar más de un archivo el tiempo de respuesta se incrementa considerablemente. Este retraso repercute en la compañía de manera negativa, en primer lugar representa un mayor costo horas-hombre y se corre el riesgo de entregar resultados erróneos, en segundo lugar la imagen

ante el cliente se ve afectada en cuanto a la calidad y rapidez del servicio de soporte técnico previamente contratado.

El requerimiento es un proceso automatizado para la manipulación de la información de tarificación de las centrales para así poder realizar el análisis de las llamadas realizadas por cada hora en un rango de tiempo mejorando el tiempo de respuesta al cliente⁴.

I.4 Las centrales en la WAN / RCD

Con el fin de garantizar que la mayor parte de los habitantes del país tengan acceso a los servicios que ofrece la RTPC y así lograr que se encuentren comunicados; las Centrales Telefónicas están distribuidas a lo largo del país y todas están interconectadas entre sí. Por otro lado, con el fin de facilitar el acceso a la información de la central⁵ por parte de los operadores, las centrales están conectadas a su vez a una WAN⁶ (Wide Area Network) (ver Figura 6). Cada central cuenta con una dirección IP⁷ que es alcanzable para los operadores a quienes por razones obvias de seguridad el personal a cargo de administrar la red les asigna una contraseña⁸.

⁴ El requerimiento es formulado por el área de Soporte Técnico de la Compañía.

⁵ Las bases de datos de los subscriptores

⁶ Red de Área Amplia a la cual se le denomina también Red Corporativa de Datos (RCD)

⁷ Internet Protocol Address

⁸ Dependiendo de los privilegios de los usuarios se define el tipo de acceso a la red.

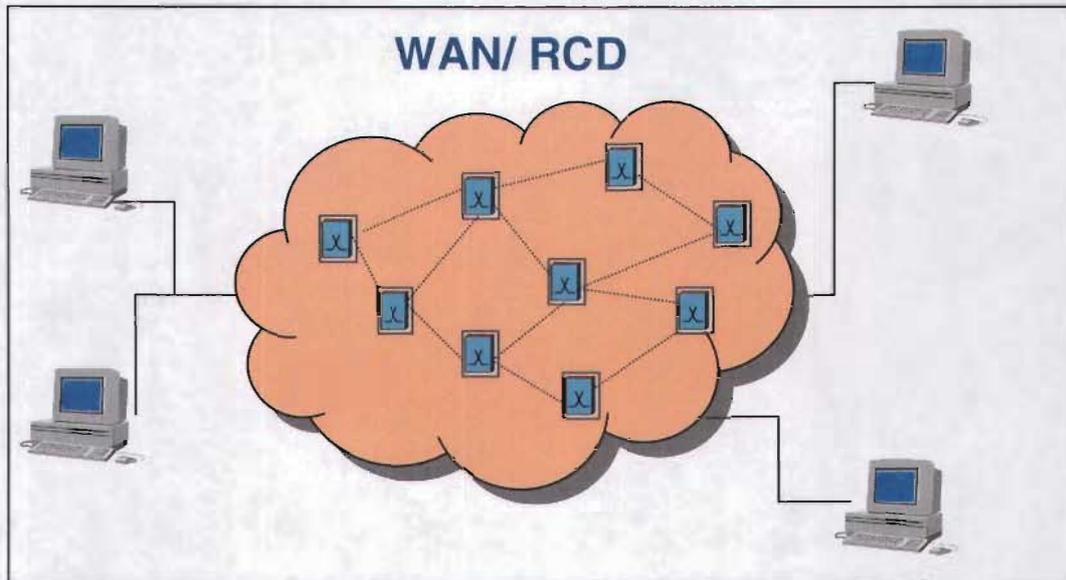


Figura 6. Red Corporativa de Datos RCD

En nuestro caso, es necesario tener acceso a dicha red a través de una computadora personal para poder recuperar las bases de datos generadas por los registros de llamadas realizadas en la central que será analizada.

Para poder realizar dicho análisis es necesario conocer el proceso de obtención de los datos y la estructura de los registros. En el siguiente capítulo se explica de forma detallada el proceso de transferencia de las bases de datos y un análisis de su contenido.

Capítulo II. Análisis de Información

La indecisión es por lo regular, peor que cometer un error.

Gerald Ford

Capítulo II. Análisis de Información

En este capítulo se presenta una secuencia de cómo los datos son generados, se describe el procedimiento de recuperación remota de los archivos de tarificación utilizando la red de datos RCD y se explica la estructura y formato de las bases de datos y su interpretación para así poder definir una estrategia de solución.

Para llevar a cabo el desarrollo de un sistema, es importante identificar los que serán nuestros *datos de entrada*, y la forma más adecuada de obtenerlos. De acuerdo al requerimiento definido en el punto 1.4, necesitamos analizar la información de tarificación, misma que se encuentra ubicada dentro de los discos magnéticos de la central en forma de archivos. Una vez obtenidos los archivos, el siguiente paso es procesarlos y entonces obtener los reportes resultantes (ver Figura 7).

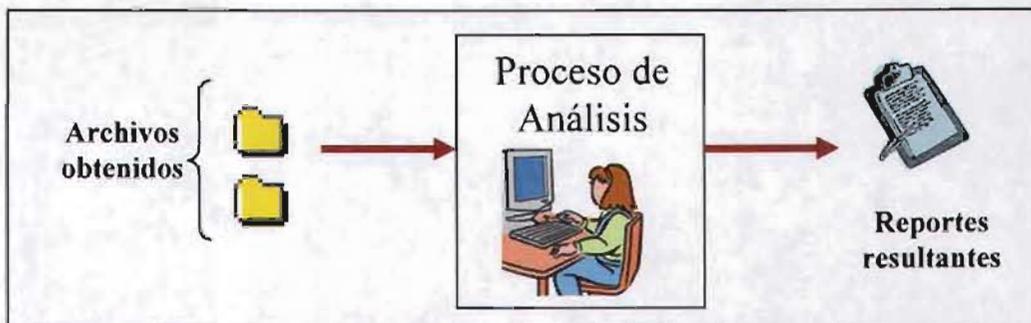


Figura 7. Proceso de desarrollo

2.1 Obtención remota de las bases de datos

Las bases de datos residen en los discos magnéticos de la central⁹, donde se guarda toda la información de la central, por ejemplo, números de directorio, prefijos, etc.; por razones de seguridad, un disco es espejo del otro, es decir, en todo momento los dos discos cuentan con la misma información.

Debido a que, como se menciona en el capítulo anterior, las centrales se encuentran ubicadas en diferentes localidades y la recopilación de los archivos de tarificación debe realizarse desde un único sitio a través de una conexión remota a la RCD.

2.1.1 Nomenclatura de los archivos

A cada archivo, el sistema propio de la central, le asigna un nombre de acuerdo a la siguiente regla:

$$\begin{array}{c} \mathbf{(C)(D)143xx.tax} \\ \\ \textit{donde} \\ \\ C,D = \textit{Disco donde se ubican} \\ \\ \textit{los archivos} \\ \\ 00 \leq xx \leq 49 \end{array}$$

⁹ Por razones de seguridad, cada una de las centrales cuenta con dos discos magnéticos definidos como "Disco C" y "Disco D".

Los archivos son generados por el sistema propio de la central en forma consecutiva a partir del 14300.tax¹⁰, y hasta llegar al 14349.tax. Después de generar el 14349.tax, se reemplazará el 14300.tax por uno nuevo, esto con el fin de no saturar la capacidad del disco y así no suspender en ningún momento la generación de los archivos. Es importante tener en cuenta que la depuración de archivos tarda aproximadamente una semana, por lo que la recuperación de éstos debe realizarse antes de que sean eliminados o reemplazados.

2.1.2 Recuperación de los archivos

La transferencia de archivos puede ser posible:

- Desde una terminal¹¹ local conectada directamente a un puerto de la central.
- Desde una terminal remota conectada a la RCD.

La transferencia de archivos desde una terminal local es mucho más rápida que la transferencia vía remota, pero es mucho más costoso en términos de tiempo y presupuesto desplazarse hasta el sitio donde se encuentra la central, por lo que es más conveniente la segunda opción: una terminal remota conectada a la RCD a través de un Módem como se muestra en la Figura 8.

Dentro del CNS se cuenta con doce computadoras y cada una tiene instalado un módem conectado a la línea telefónica, como lo señala Morguitsu (Año 2000, p.662) cuando dos módems se conectan, es posible

¹⁰ La extensión “.tax” indica que se trata de un archivo de tarificación. es la abreviación del término “taxation” en inglés.

¹¹ Se utiliza una Computadora Personal como terminal y un puerto RS232.

establecer una comunicación efectiva entre dos sistemas host remotos, esta comunicación permite que se puedan compartir archivos e impresoras y la ejecución de aplicaciones remotas. Un nombre de usuario y una contraseña de sólo lectura son proporcionados por el personal encargado de administrar la RCD con el fin de que un usuario del CNS logre conectarse y así pueda entonces llevar a cabo la transferencia de los datos.

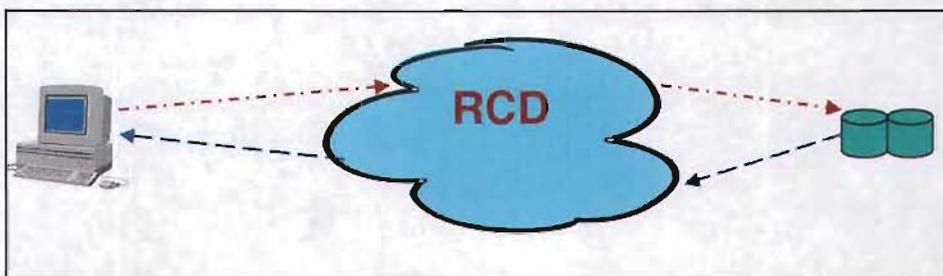


Figura 8. Recuperación remota de datos

La RCD es una red que trabaja con un sistema UNIX y utiliza ruteadores para incluir sistemas remotos, así como el paquete de protocolos TCP/IP que se encuentra en el corazón del tratamiento de redes en los sistemas UNIX.

Cada central cuenta con una dirección IP dentro de la red, y con un servidor FTP¹². En promedio el tamaño de los archivos varía entre 12 y 16 Megabytes, por esta razón la transferencia es lenta y puede variar entre 2 y 5 horas.

¹² File Transfer Protocol. Proporciona servicios de transferencia de archivos.

La aplicación del protocolo de transferencia de archivos FTP se puede utilizar para transferir archivos entre hosts. FTP puede interactuar desde la línea de comandos. Windows proporciona una implementación estándar del programa FTP para transferir archivos y está disponible desde un programa DOS. Para obtener los archivos remotamente utilizaremos la aplicación FTP.

Se abre una sesión FTP desde la ventana *command*, con la dirección IP que corresponde a la central a analizar, tomemos como ejemplo los datos de la central "**Apodaca**" que se encuentra ubicada en la ciudad de Monterrey, Nuevo León y su dirección IP es "13.138.55.225". Se aplica el procedimiento siguiente:

```
C:\Taxation>ftp
ftp>open 13.20.35.227
Connected to 13.138.55.225
220 ftp server (Monday, Feb 23 12:12:34 EST 2004) ready.
.
user: oper1
331 password required for oper1.
Password: *****
user oper1 logged in
ftp>_
```

Los datos que están marcados en letras negritas, son los que introducirá el operador. A partir de este momento ya establecimos la sesión en la central Apodaca y podemos recuperar los archivos de nuestro interés.

Primero que nada, se verifican la fecha y hora en que los archivos fueron creados para transferir el que nos servirá para nuestro análisis posterior. Como ejemplo, recuperaremos la información de la central correspondiente al día 22 de febrero hasta antes del medio día. Para ver la fecha de creación de los archivos utilizamos el comando `ls -l`, como se muestra a continuación:

```
ftp>ls -l
-rwxr-x 1 root system 12288000 Feb 21 09:15p c14303.tax
-rwxr-x 1 root system 12288000 Feb 21 09:15p d14303.tax
-rwxr-x 1 root system 7252740 Feb 22 02:10a c14304.tax
-rwxr-x 1 root system 7252740 Feb 22 02:10a d14304.tax
-rwxr-x 1 root system 14204300 Feb 22 08:00a c14305.tax
-rwxr-x 1 root system 14204300 Feb 22 08:00a d14305.tax
-rwxr-x 1 root system 12288000 Feb 22 02:27p c14306.tax
-rwxr-x 1 root system 12288000 Feb 22 02:27p d14306.tax
. . .
200 PORT command successful.
226 ASCII transfer complete
ftp>_
```

Como podemos observar, el archivo más adecuado para nuestro análisis es el "c14306.tax" ya que fue cerrado a las 2:27 p.m. y contiene registros a partir de las 8:00 a.m., ya identificado procedemos a transferirlo.

Para la transferencia de los archivos se utilizan comúnmente los comandos **get** y **mget**¹³, como sólo requerimos un archivo, utilizamos el comando **get**. Pero antes debemos indicar el tipo de transferencia, **bin** para archivos en formato binario, el comando **hash** se utiliza para poder ver en pantalla el progreso de la transferencia, lo cual es conveniente ya que por el tamaño del archivo puede durar entre 3 y 4 horas en promedio.

¹³ **get** para transferir un solo archivo y **mget** para transferir múltiples archivos.

2.2 Formato y estructura de los archivos

Cada vez que un abonado realiza una llamada telefónica, ésta es almacenada en los discos de la central en el momento del cuelgue. Los datos generados por las llamadas realizadas son almacenados de manera que cada llamada forma un registro y a su vez un conjunto de registros forma un archivo.

La base de datos está formada por un número finito de archivos que se generan de manera dinámica. Las bases de datos fueron diseñadas de manera tal, que a simple vista no es posible entender e interpretar los registros almacenados. Las centrales **no** cuentan con un manejador común de bases de datos (Oracle, SQL, Access, etc.). Toda la información es almacenada en archivos con formato "binario" los cuales tienen una estructura definida, de manera que puedan ubicarse los registros y campos de tipo y longitud determinados, estos archivos son llamados también CDF (Call Data Record Files).

El hecho de que los archivos tengan un formato binario es una ventaja ya que el acceso binario nos da un control completo sobre el archivo, debido a que los bytes en el archivo pueden representar cualquier dato, pero sobre todo permiten conservar un tamaño más pequeño que un archivo de texto.

2.2.1 Estructura de los archivos CDF

La estructura de los archivos CDF es la siguiente:

Definition of CDF-File:			
CDF-File	::=	SEQUENCE {	
header	[0]	FileHeaderRecord	OPTIONAL,
body	[1]	SEQUENCE OF RecordContent,	
trailer	[2]	Trailer	OPTIONAL
		}	

Al inicio del archivo se encuentra el "header" (encabezado), en el cual se almacena información sobre la hora y fecha en que se creó el archivo, tamaño de éste, el protocolo utilizado para su transferencia, nombre de la central, tipo, etc. El header tiene una longitud de 2048 bytes.

En la parte "body" (Sequence of RecordContent) están almacenados los registros de llamadas de manera secuencial, cada llamada genera un registro, el cual es almacenado inmediatamente después de que se generó el anterior. Es importante hacer notar que cada registro se almacenará inmediatamente después de que se realiza la interrupción de la llamada por parte del abonado llamante, es decir, al cuelgue del teléfono. Cada registro mide 110 bytes.

Por último, al final del archivo tenemos el "trailer" en donde se registra información de control como por ejemplo, último número en la secuencia, contadores, sumatorias, etc.

En promedio cada archivo contiene alrededor de 100,000 registros y se generan entre 4 y 6 durante el transcurso de un día. La central tiene un sistema de mantenimiento para que los archivos se cierren al transcurrir un lapso de tiempo determinado para evitar que éstos crezcan demasiado y así facilitar el proceso de transferencia vía remota. El tamaño en promedio de cada archivo es de 11,600 Kilobytes, y puede variar de acuerdo a la carga de tráfico que se registre en la central, en ocasiones los archivos llegan a tener un tamaño de hasta 16 Megabytes.

2.2.2 Formato de los archivos CDF

Los archivos CDF residen en los discos de la central en formato "binario", una vez recuperados es posible editarlos utilizando un editor como "HexWorkShop", mismo que puede ser extraído de Internet gratuitamente¹⁴ y que son útiles para operaciones de verificación de datos al momento de realizar el programa.

2.3 Los registros CDR

En el momento del descuelgue del teléfono por parte del abonado llamante, se validan sus datos, una vez detectado que se trata de una línea telefónica activa, se le otorga el tono de marcado. En ese momento se abre un registro con los datos del abonado llamante: número de directorio (DN¹⁵), tipo de línea, etc. Si no se logra establecer la llamada, el registro no se guarda, en caso contrario se inicia un conteo de duración de llamada, y al momento del cuelgue por parte del abonado llamante, se almacena el registro conteniendo un número consecutivo, además de los detalles relevantes acerca de la llamada: hora de inicio de llamada, hora de término, número

¹⁴ Página Web: www.hexworkshop.com

¹⁵ DN = Directory Number

destino, etc. Los registros de llamadas son denominados CDR (Call Data Record).

Es importante mencionar que, los datos en el archivo se encuentran ordenados en forma ascendente, solamente con respecto al número de secuencia (Record Sequence Number), el cual no está relacionado con la hora de la llamada, es decir, los registros no llevan un orden cronológico dentro del archivo.

2.3.1 Estructura de los registros CDR

La tabla que se muestra en la siguiente página representa la estructura de un registro de llamada CDR.

Field ¹⁶	Identification	Bytes ¹⁷	Position	Value
1	Service Code	1	1	HEX
2	Record Length	2	2	HEX
3	Output Cause	1	4	BCD
4	Register Number	1	5	HEX
5	Call Identification Number	4	6	HEX
6	Record Sequential Number	3	10	HEX
7	Calling Number	9	13	BCD
8	Call Tax Category	1	22	HEX
9	Called Number	9	23	BCD
10	Called (B) Party Status	1	32	HEX
11	Reason for Disconnection	1	33	HEX
12	Fault Code of the Incoming Part	2	34	HEX
13	Date of the Beginning Charging (yy mm dd)	3	36	BCD
14	Time of the Beginning Charging (hh mm ss)	3	39	BCD
15	Duration Time of Charging	3	42	BCD
16	Duration Time of Re-answer Supervision	2	45	HEX
17	Duration Time from Record Started till the Beginning of Charging	2	47	HEX
18	Selected Carrier	3	49	BCD
19	Outgoing Route Mnemonic	7	52	ALPH
20	Incoming Route Mnemonic	7	59	ALPH
21	Rerouting Indicator	1	66	BCD
22	Status of the call	1	67	HEX
23	Indicator of Normal or Abnormal Release	1	68	BCD
24	Outgoing Route Index	2	69	BCD
25	Incoming Route Index	2	71	BCD
26	Supplementary Service	1	73	HEX
27	Abbreviated Number	2	74	BCD
28	Related Call	4	76	HEX
29	Correct or Incorrect Dialling Indication	1	80	BCD
30	Facility Action	1	81	HEX
31	Subscriber X Directory Number	9	82	BCD
32	Charged Party	1	91	HEX
33	Immediate Service Indicator	1	92	BCD
34	Price	4	93	BCD
35	Origin of Charging	2	97	HEX
36	Tariff Class	1	99	HEX
37	Tariff Switch Over Indication	1	100	BCD
38	Intelligent Network Service Indication	10	101	HEX
	Total Record Size	110		

¹⁶ Campos definidos en un registro CDR. Documentación interna.

¹⁷ Tamaño del campo dado en bytes

HEX – Formato Hexadecimal

BCD – Formato Binary Coded Decimal

Como se observa en la tabla anterior, el tamaño de cada registro es de 110 bytes y se manejan dos tipos de datos que se describen a continuación.

- a. **HEX.** El dato se encuentra almacenado en sistema hexadecimal, es necesario realizar una operación de conversión a sistema decimal para así conocer el valor del campo. Por ejemplo, en el caso del campo número 2 (Record Length) que nos da la longitud del registro, es un dato constante, ya que todos los registros en todos los archivos tienen el mismo tamaño. Tomando los primeros dos bytes del registro de tarificación, tenemos que:

Pos	1	2	3	4	5	6
	01	00	6E	FF	01	99 . . .
	Campo 2. 2 bytes = 006E					
	⇒ Record Length = HEX(006E) = 110					

El valor decimal del campo Record Length es igual a 110. La longitud del registro es, entonces 110 bytes, misma que se mantendrá constante para todos los registros.

- b. **BCD.** El dato se encuentra almacenado en sistema decimal, no es necesario realizar una operación conversión para conocer el valor del campo. Por ejemplo, en el caso del campo número 7 (Calling

Number) que nos da el DN del abonado llamante¹⁸, se interpreta de la siguiente manera:

Pos	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
	01	00	6E	FF	01	99	10	D6	45	A0	46	4F	81	83	15	07	35	BB	BB	BB	BB	.
																		Campo 7 contiene 9 bytes = 8183150735BBBBBBBB				
																		⇒ Calling Number = 8183150735BBBBBBBB				

El valor del campo Calling Number es igual a 8183150735BBBBBBBB. Aunque en la práctica y para el caso específico de México, sólo se toman los primeros 5 bytes, ya que los últimos 4 están reservados para el caso de un cambio en el número de dígitos del sistema de marcación. Por lo tanto, el DN del abonado llamante es: 8183150735.

2.3.2 Interpretación de los registros CDR

Los registros CDR contienen un total de 38 campos, pero para el proceso de análisis sólo vamos a tomar los que son de nuestro interés, los demás no son requeridos en nuestro proceso y algunos de ellos no tienen aplicación ni siquiera para el mercado de telecomunicaciones en México. Los campos que se requieren para nuestro análisis son:

¹⁸ El DN del abonado llamante es el número telefónico del abonado que realiza la marcación y a quien se le hará el cargo por la llamada.

Field	Identification	Bytes	Position	Value
6	Record Sequential Number	3	10	HEX
13	Date of the Begining Charging	3	36	BCD
14	Time of the Begining Charging	3	39	BCD
	Total Record Size	110		

¿Por qué son necesarios sólo estos tres campos? Recordemos que en el punto 1.3 donde se plantea el requerimiento, se explica que la necesidad de verificar el número de llamadas realizadas por hora. La fecha y hora de llamada son básicas (campos 13 y 14), y por supuesto el número de secuencia (campo 6) nos indicará si hubo algún salto en la grabación de los registros.

Como ejemplo analizaremos CDRs de un archivo de la central Apodaca:

```

01 00 6E 00 01 87 80 28 45 A0 46 60 81 83 46 41 34 BB BB BB BB 02 81 83 68 14 00 BB
BB BB BB 00 09 FF FF 04 02 22 08 15 40 00 01 55 00 00 00 0B FF FF FF 30 30 32 5F 30 31
32 4C 4F 43 41 4C 20 20 00 00 00 01 2B FF FB 0C FF FF FF FF FF 00 08 FF FF FF FF
FF FF FF FF FF 00 FF FF
01 00 6E 00 01 8B A0 29 C2 A0 46 61 81 83 46 80 84 BB BB BB BB 02 82 12 12 05 22 BB
BB BB BB 00 09 FF FF 04 02 22 08 17 30 00 05 03 00 00 00 0D 01 12 3B 30 30 32 5F 30 31
32 4C 4F 43 41 4C 20 20 00 00 00 01 2B FF FB FF FF FF FF FF FF FF FF 00 FF FF FF
FF FF FF FF FF 00 FF FF
01 00 6E FF 01 7C 10 08 34 A0 46 62 81 83 33 11 53 BB BB BB BB 02 81 83 36 21 88 BB
BB BB BB 23 05 FF FF 03 04 02 22 08 17 50 00 00 00 00 00 0B FF FF FF 30 30 32 5F 30 31
32 4C 4F 43 41 4C 20 20 00 01 00 01 2B FF FB 0C FF FF FF FF FF FF 00 08 FF FF FF FF
FF FF FF FF FF 00 FF FF
    
```

Si comprimimos el registro y sólo mostramos los campos de nuestro interés, tenemos:

```

A0 46 60 04 02 22 08 15 40
A0 46 61 04 02 22 08 17 30
A0 46 62 04 02 22 08 17 50
    
```

Donde los primeros 3 bytes nos indican el campo del número de secuencia del registro en el archivo (Record Sequential Number), los siguientes 3 bytes corresponden al campo de la fecha en que se inició la llamada (Date of the Beginnig Charging), y los últimos 3 bytes del campo nos indican la hora de inicio de llamada (Time of the Beginnig Charging).

Luego entonces, el primer registro se interpreta de la siguiente manera:

Record Sequential Number: A0 46 60, es un dato con valor HEX, por lo que al realizar la conversión al sistema decimal tenemos que el número de secuencia es:

HEX(A0 46 60)=10503776

Date of the Beginnig Charging: 04 02 22

que es de tipo BCD, por lo que la fecha de inicio de llamada es:

yy mm dd

04 02 22 = 2004 02 22, 22 de febrero del 2004

Time of the Beginnig Charging: 08 15 40 también de tipo BCD, por lo que la hora de inicio de llamada es:

hh mm ss

08 15 40 = 08:15:40, 8 horas, 15 minutos, 40 segundos

Por lo tanto, el abonado llamante inició su llamada el día 22 de febrero del 2004 a las 8:15:40.

2.4 Información de salida

Los datos de salida son en realidad los reportes esperados después de completar el proceso de análisis. Los reportes deben contener los siguientes datos:

Para el Desglose:

Encabezado:

Nombre de la Central.

Fecha de emisión reporte

Cuerpo (para cada archivo):

Nombre del archivo, fecha de llamada, reporte de llamadas por cada hora, subtotal

Al final:

Total de llamadas por día, por hora, Gran Total

El reporte debe lucir como se muestra en la Figura 9, en el entendido de que deben incluirse los datos de llamadas registradas en el archivo desde las 0:00 horas hasta las 23:00.

Nombre de Central:															
Fecha de Emisión:															
Archivo	Fecha	HORA													Total
		0:00	1:00	2:00	3:00	4:00	5:00	6:00	...	19:00	20:00	21:00	22:00	23:00	
Resumen															
	Fecha								...						Total
Gran Total															

Figura 9. Formato del reporte

En los campos correspondientes a la hora, se registraran las llamadas ocurridas entre el inicio y término de la hora indicada, por ejemplo, para las 0:00 se registran llamadas desde las 0 horas 0 minutos 0 segundos y hasta las 0 horas 59 minutos 59 segundos y así sucesivamente para las demás horas.

También desea conocerse si hubo pérdida en la secuencia, en tal caso es necesario indicar en los números donde se perdió la secuencia.

Para la Secuencia:

Inicio de proceso. hh:mm:ss

Número inicial en la secuencia

.....

Número donde se presentó un salto en la secuencia.

.....

Número final en la secuencia

Fin de proceso. hh:mm:ss

Resultado del análisis

El proceso de generación de reportes tomando como base los archivos de tarificación realizado manualmente implica demasiado tiempo. Detectar una pérdida en la secuencia es un proceso tardado y tedioso, mas allá de la gran cantidad de registros existentes en cada archivo, la interpretación de los datos requiere un poco más de concentración.

Hasta este momento tenemos identificado el problema y los resultados que se esperan. En el siguiente capítulo se define una propuesta para hacer más eficiente la obtención de los reportes reduciendo tiempo y esfuerzo a los operadores, que son los encargados de entregar los resultados del análisis.

Capítulo III. Propuesta de Solución

Si quieres ver un arco-iris, tendrás que soportar primero la lluvia.

Dolly Parton

Capítulo III. Propuesta de Solución

En este capítulo se presenta una propuesta de solución que involucra el desarrollo de un sistema, así como los algoritmos, procesos principales y la implementación.

Los algoritmos de los procesos principales se expresan en forma de diagramas de flujo, tomando en cuenta que como señala Gerez (1998) los diagramas de flujo son una forma universal para expresar la estrategia de solución de un problema en una computadora digital.

3.1 Diagnóstico de las circunstancias

3.1.1 Identificación de recursos

En primera instancia deben ser identificados los recursos con los que se cuentan para poder llevar a cabo el desarrollo del sistema. El CNS que es el departamento encargado de generar los reportes cuenta con 12 computadoras personales que tienen las siguientes características:

Hardware:

- Procesador Pentium IV,
- Memoria RAM de 256 MB
- Disco duro de 40 GB,

- Módem → conexión a RCD

Software:

- Windows 2000
- Office 2000

El sistema desarrollado deberá ser instalado y operado en cada una de las PCs.

Para la parte del desarrollo contamos con una computadora que además de tener las mismas características tanto de Hardware como de Software, tiene instalado el paquete Visual Studio versión 6.0 Professional Edition, e incluye herramientas de desarrollo como: Visual C++, Visual J++, Visual Basic, Visual InterDev, Visual FoxPro y Visual SourceSafe¹⁹.

3.1.2 Solución propuesta

Se propone en primera instancia el desarrollo e implementación de un sistema para optimizar el proceso de análisis de los registros de tarificación, utilizando como herramienta de programación Visual Basic 6.0.

El sistema deberá ser capaz de:

- Operar de manera independiente en cada una de las PCs descritas en el punto **3.1.1 Identificación de recursos**.

¹⁹ MSDN Library Visual Studio 6.0, Visual Studio Product Documentation.

- Obtener los archivos CDFs de manera remota, esto es, transferirlos desde la CT hacia la PC.
- Realizar el procesamiento de éstos.
- Enviarlos a una hoja de cálculo para su correcta interpretación²⁰.
- Proporcionar una interfaz gráfica y amigable al usuario.

Al sistema se le asignó el nombre de "Tax Converter".

3.1.3 Soluciones alternativas

El diseño de un sistema es definitivamente la solución más viable para optimizar el tiempo de procesamiento de los archivos, pero es importante definir el lenguaje de programación a utilizar para llevar a cabo el desarrollo de dicho sistema.

Una de las premisas es utilizar los recursos con los que cuenta la compañía, como se menciona en el punto 3.1.1, el Software para desarrollo con el que se cuenta es el "Visual Studio versión 6.0 Release", que involucra la familia de los siguientes lenguajes de programación:

- Visual C++
- Visual J++
- Visual Basic

²⁰ Se utilizará Microsoft Excel contenido en el paquete Microsoft Office 2000.

3.1.4 Justificación

Procedemos a analizar cual de los tres lenguajes de programación con los que se cuenta es la opción más viable:

a) Visual C++:

Es un lenguaje de propósito general conocido por su eficiencia, bajo costo y portabilidad. Es una buena opción para casi cualquier clase de programa. Se considera una excelente opción por la flexibilidad con la que cuenta para la manipulación de archivos.

b) Visual J++:

Es una herramienta integrada para Windows, con programación Java que involucra el objeto Dynamic HTML implementado en el Internet Explorer, el cual permite manipular dinámicamente documentos HTML.

Para el desarrollo de nuestro sistema no es la mejor opción ya que aunque poderoso, el lenguaje es más bien orientado a diseño de páginas Web.

c) Visual Basic:

Es una herramienta que proporciona una manera fácil y rápida de crear aplicaciones para Windows, la programación en Visual Basic

permite crear aplicaciones atractivas y útiles explotando al máximo la Interfase Gráfica de Usuario (GUI). En lugar de escribir un gran número de líneas de código para definir la apariencia y localización de los elementos gráficos, simplemente se incrustan en la pantalla objetos pre-diseñados, a los cuales se les definen propiedades específicas de apariencia, localización, etc.

Aunado a la facilidad para crear la interfase gráfica, Visual Basic nos ofrece también la ventaja de poder manipular archivos de manera eficaz y confiable. Al compararlo con Visual C++, el tiempo de respuesta para acceso lectura/escritura a los archivos binarios no tuvo variación alguna.

Otro punto importante a considerar es que los reportes deben ser enviados a una hoja de cálculo Excel, cualquiera de los lenguajes mencionados nos brinda la posibilidad de manipular hojas de cálculo utilizando las librerías adecuadas.

Se tomó la decisión de utilizar Visual Basic, por las siguientes razones:

- La facilidad que ofrece para el diseño de la interfase gráfica
- La rapidez con que manipula los archivos
- La integración de las librerías de Excel²¹ facilitan la creación y manipulación de los reportes.

²¹ Se utilizó la librería "Microsoft Excel 8.0 Object Library" incluida en el paquete Office 2000.

3.2 Descripción de procedimientos.

Los principales procedimientos para llevar a cabo el análisis de los archivos de tarificación son los que se describen a continuación:

1. Obtención de los archivos de tarificación
2. Procesamiento de archivos
3. Envío de reportes resultantes a Excel

3.2.1 Obtención de los archivos de tarificación

Como se describe en el punto **2.1.2 Recuperación de los archivos** los archivos de tarificación pueden obtenerse utilizando el protocolo de transferencia FTP abriendo una ventana command y proporcionando los comandos de manera manual.

El usuario deberá tener a la mano los siguientes datos para transferir el archivo:

- Nombre de usuario y password de la RCD²².
- La dirección IP de la CT donde se encuentra el archivo a analizar.
- El nombre del archivo.
- La ruta donde desea guardar el archivo recuperado.

La Figura 10 muestra el procedimiento a seguir en un diagrama de flujo para lograr la obtención del CDF.

²² El usuario debe tener previamente establecida una sesión en la Red Corporativo de datos.

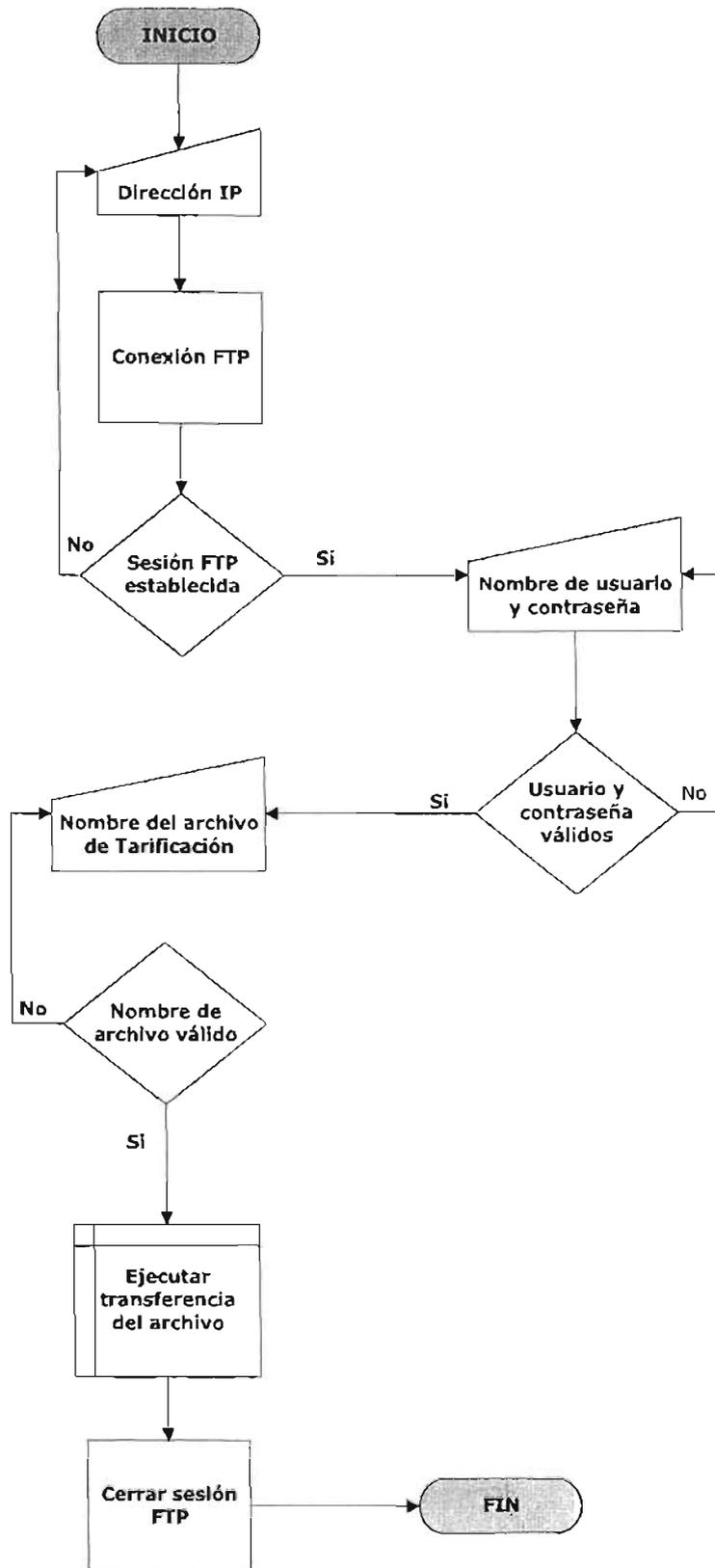


Figura 10. Diagrama de flujo: Transferencia de archivos

3.2.2 Procesamiento de archivos

Una vez recuperados los archivos de tarificación es posible llevar a cabo el proceso de análisis, la utilización del sistema parte del momento en que los archivos hayan sido recuperados como se describe en el punto anterior.

El análisis debe realizarse de manera automática cuando el usuario así lo indique. Se requieren realizar dos procesos esenciales:

1. El desglose de llamadas por fecha y hora
2. Verificar la secuencia de los registros

Tomamos como ejemplo un fragmento obtenido de un CDF presentado el punto 2.3.2.

```

01 00 6E 00 01 87 80 28 45 A0 46 60 81 83 46 41 34 BB BB BB BB 02 81 83 68 14 00 BB
BB BB BB 00 09 FF FF 04 02 22 08 15 40 00 01 55 00 00 00 0B FF FF FF 30 30 32 5F 30 31
32 4C 4F 43 41 4C 20 20 00 00 00 01 2B FF FB 0C FF FF FF FF FF FF 00 08 FF FF FF FF
FF FF FF FF FF 00 FF FF
01 00 6E 00 01 8B A0 29 C1 A0 46 61 81 83 46 80 84 BB BB BB BB 02 82 12 12 05 22 BB
BB BB BB 00 09 FF FF 04 02 22 08 17 30 00 05 03 00 00 00 0D 01 12 3B 30 30 32 5F 30 31
32 4C 4F 43 41 4C 20 20 00 00 00 01 2B FF FB FF FF
FF FF FF FF FF 00 FF FF
01 00 6E FF 01 7C 10 08 34 A0 46 62 81 83 33 11 53 BB BB BB BB 02 81 83 36 21 88 BB
BB BB BB 23 05 FF FF 03 04 02 22 08 17 50 00 00 00 00 0B FF FF FF 30 30 32 5F 30 31
32 4C 4F 43 41 4C 20 20 00 01 00 01 2B FF FB 0C FF FF FF FF FF FF FF 00 08 FF FF FF FF
FF FF FF FF FF 00 FF FF

```

Donde podemos ver que los datos que nos interesan para nuestro análisis se encuentran localizados en las siguientes posiciones para cada CDR:

	Posición	Tamaño (bytes)
Número de secuencia	10	3
Fecha	36	3
Hora	39	3

También recordemos que cada CDR cuenta con un total de 110 bytes de longitud.

3.2.2.1 Desglose de llamadas

A continuación se describe el proceso a seguir para llevar a cabo el desglose de llamadas. El diagrama de flujo contenido en la Figura 11 representa gráficamente el proceso:

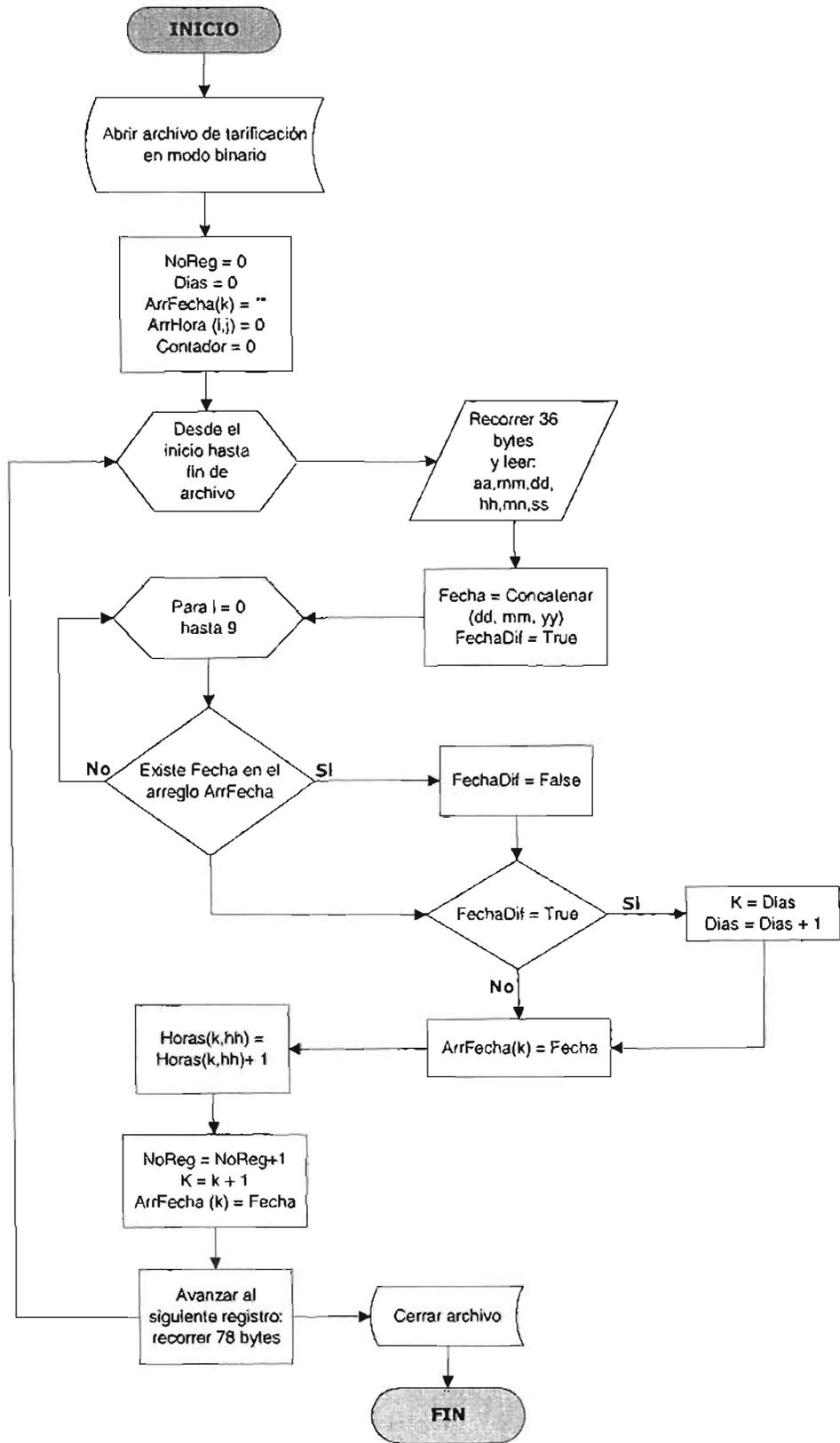


Figura 11. Diagrama de Flujo 'Desglose de llamadas'

Como resultado obtenemos los arreglos ArrFecha (unidimensional) y ArrHora (bidimensional), estructurados con el siguiente formato:

ArrFecha

K	Fecha
0	
1	
2	
3	
4	
5	
6	
7	
8	
9	

ArrHora

Hora / Dia	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	..	22	23
0																		
1																		
2																		
3																		
4																		
5																		
6																		
7																		
8																		
9																		

3.2.2.2 Envío de datos a Excel

La información que se requiere analizar, es la que resulta del proceso descrito en el punto 3.2.2.1 "Desglose de llamadas". El reporte será generado en base a los datos obtenidos en los arreglos "ArrFecha" y "ArrHora".

Visual Studio puede hacer uso de una librería²³ para poder enviar datos a una hoja de cálculo de "Excel" (ver Figura 12). Ésta se utilizará para que la creación del reporte sea totalmente automática.

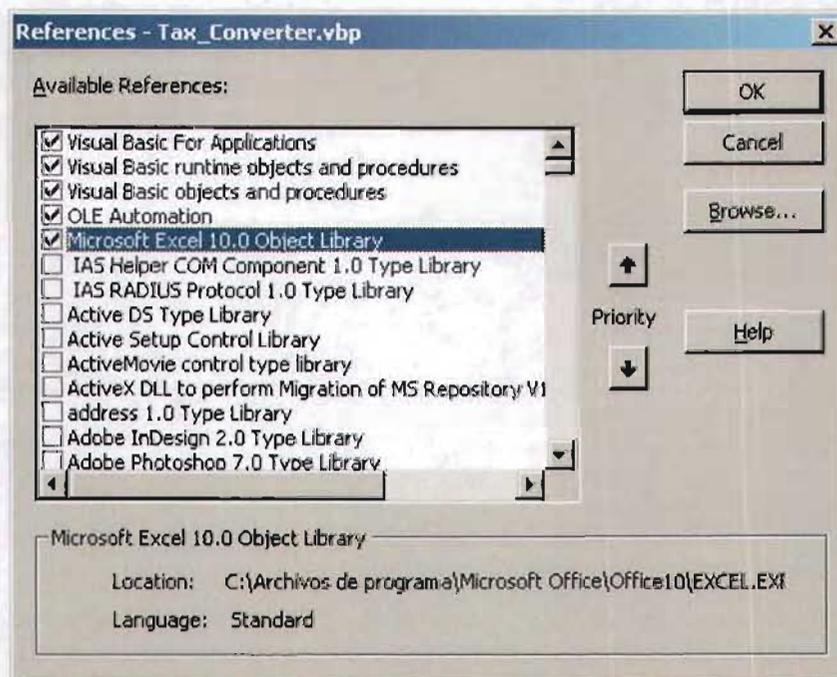


Figura 12. Visual Basic References

²³ Microsoft Excel 10.0 Object Library

Una vez integrada la librería tenemos la posibilidad de generar un nuevo libro de Excel, vaciar en éste los datos del arreglo obtenido y darle formato al archivo. Se describe a continuación el procedimiento en un diagrama de flujo (Figura 13).

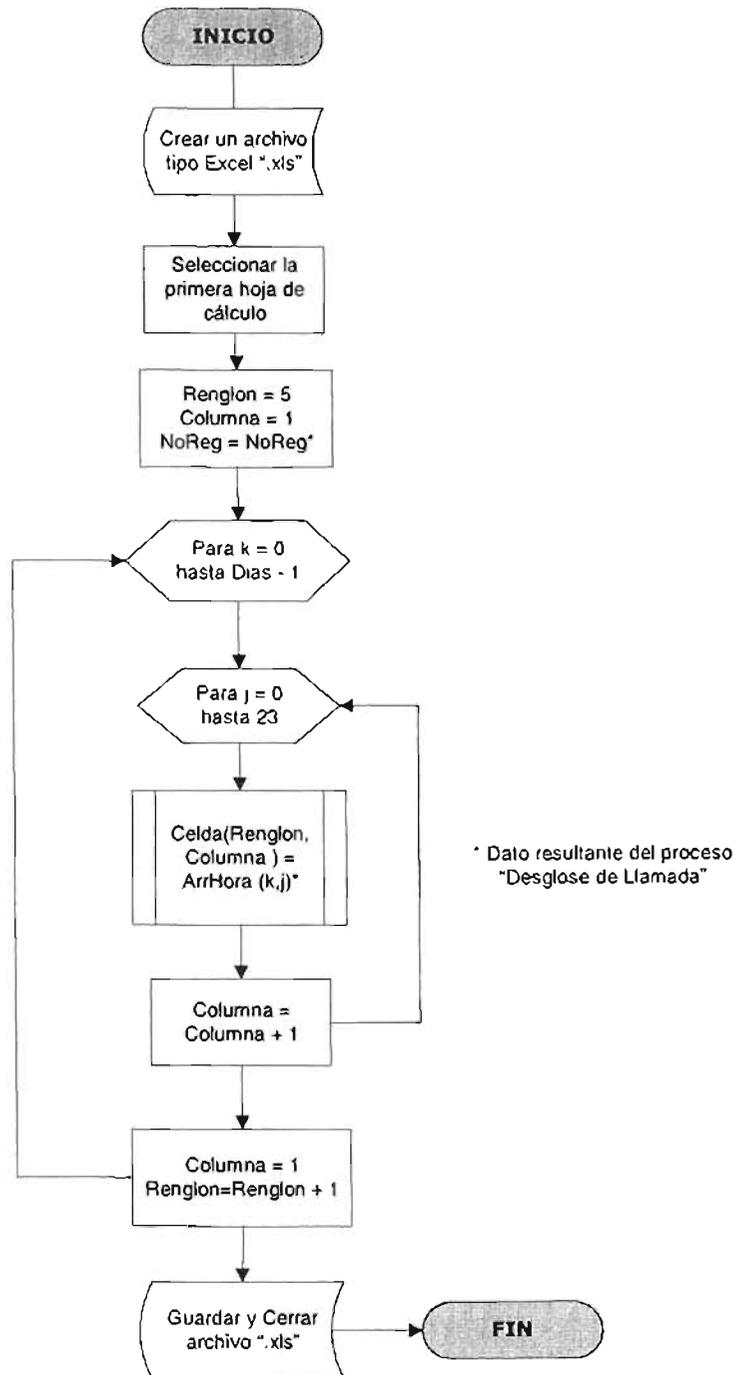


Figura 13. Diagrama de Flujo: Envío de resultados a Excel

3.2.2.3 Verificación de secuencia

El proceso de 'verificación de secuencia' se realiza para identificar si existe un salto en la numeración de los CDR's, lo que indicaría una pérdida de registros de llamadas en el archivo. A continuación la Figura 14 describe la lógica a seguir para verificar esta secuencia en un diagrama de flujo :

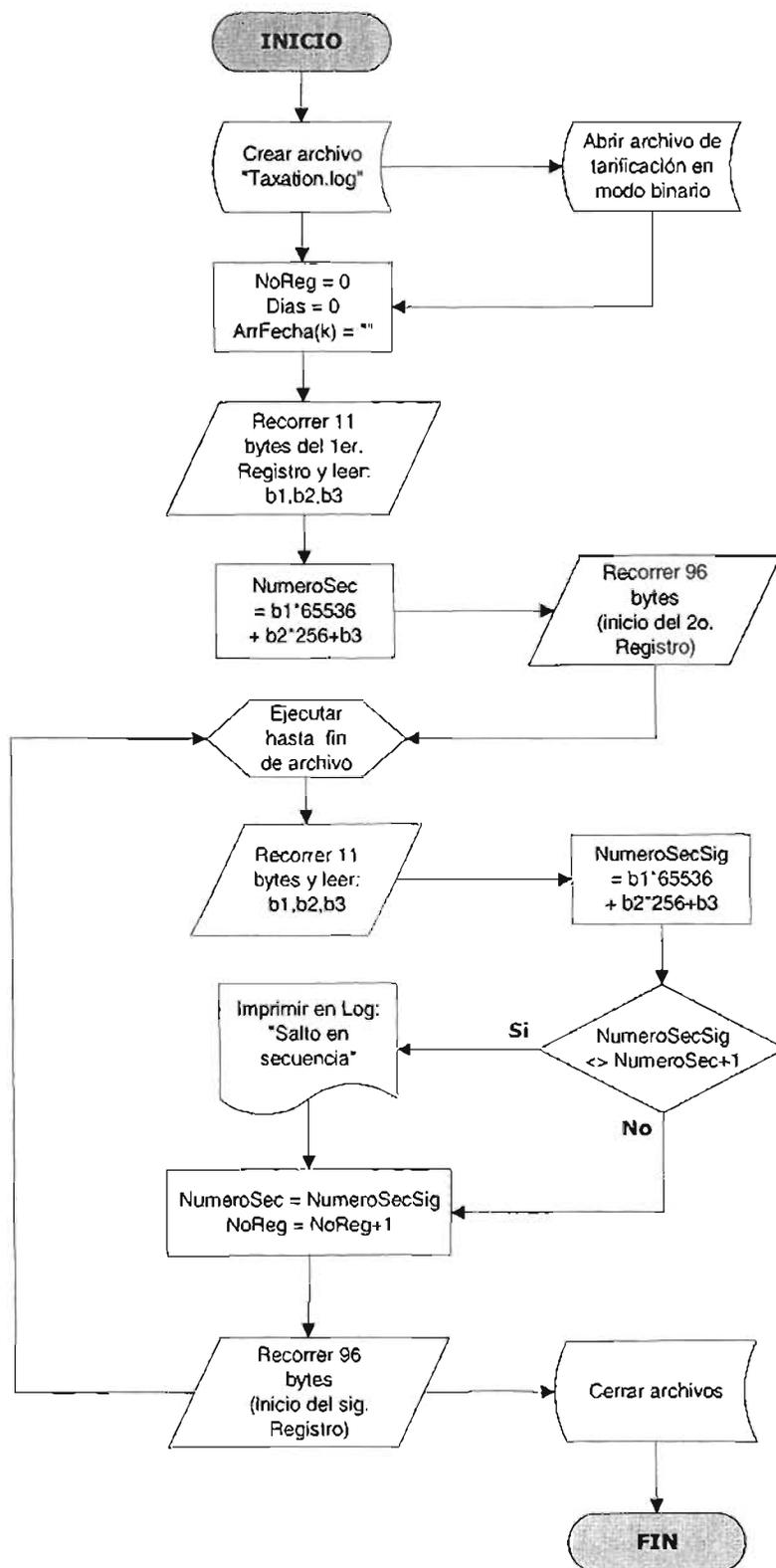


Figura 14. Diagrama de Flujo 'Verificación de Secuencia'

3.3 La interfase gráfica

La interfase gráfica de usuario (GUI) permite el manejo directo de la representación gráfica en pantalla utilizando el teclado o el mouse de acuerdo a la definición de Kendal (1997, p. 665).

Para nuestro sistema la interfase gráfica se diseñó creando una ventana de diálogo a partir de la cual el usuario eligirá la ruta de los archivos a procesar, y dará la orden al para que se ejecute el proceso deseado. Esta ventana de diálogo se crea utilizando objetos de Visual Basic y luce en primera instancia como se muestra en la Figura 15.

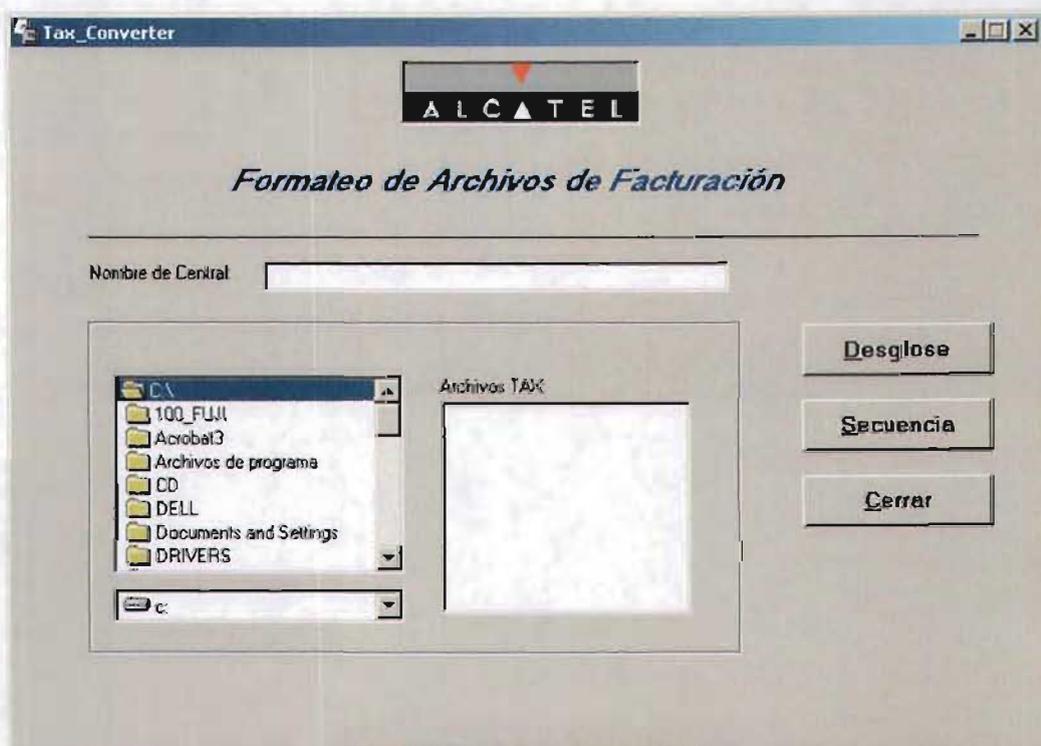
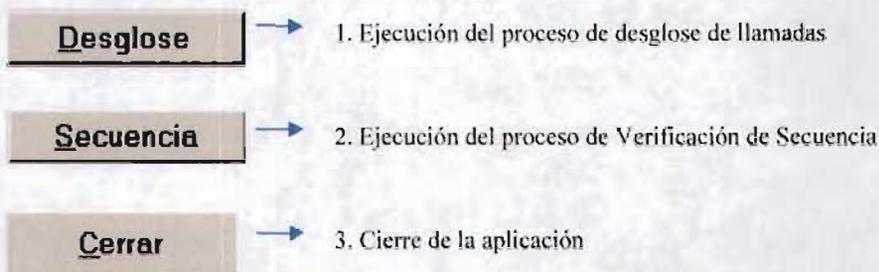


Figura 15. Interfase gráfica de usuario

Los procesos se ejecutan a petición del usuario dando click en cada uno de los "botones de comando" como se describe:



Para los botones de comando 1 y 2 es necesario seleccionar el(los) archivo(s) a procesar, (el sistema está diseñado para procesar 1 o más archivos a la vez), en este punto, contamos con que el archivo de tarificación ya fue transferido de los discos de la central a la PC donde se ejecutarán dichos procesos.

En caso de no tener previamente seleccionado el archivo a procesar, el sistema no podrá realizar ningún proceso y mandará un mensaje indicando que no puede continuar como se muestra en la Figura 16:



Figura 16. Mensaje

Cuando el (los) archivos han sido seleccionados el sistema iniciará el proceso mostrando el porcentaje de avance como se muestra en la Figura 17.

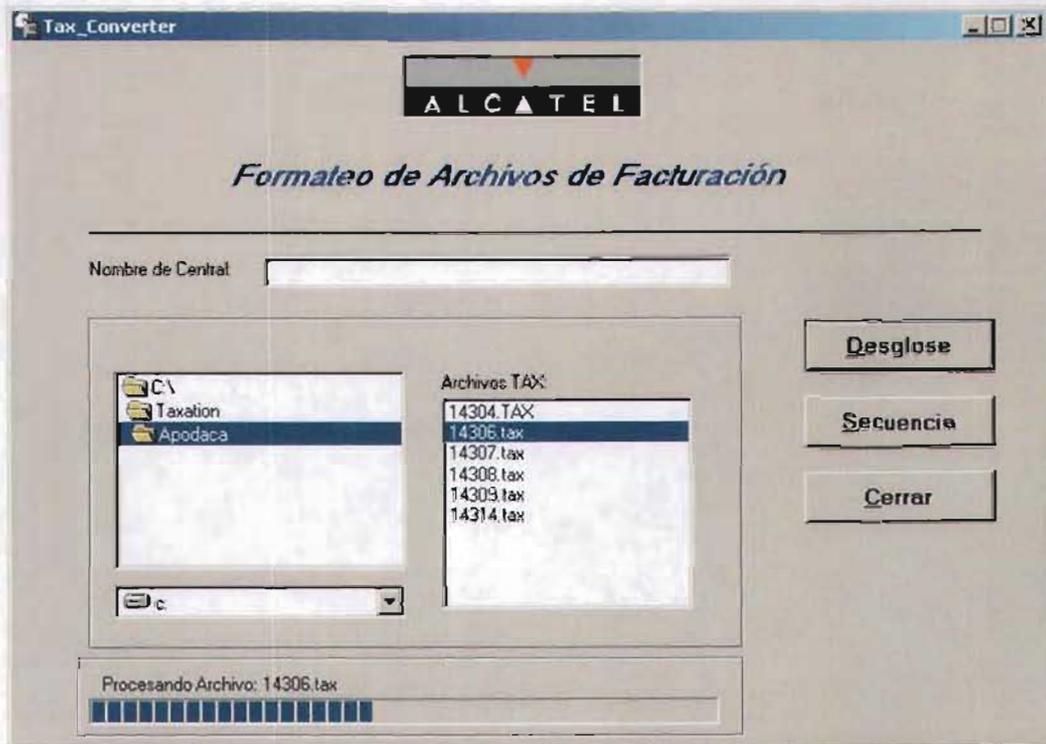


Figura 17. Avance de Procesamiento

La interfase gráfica que se presenta facilita al usuario la ejecución de los procesos descritos de una manera amigable, así mismo el tiempo de procesamiento de archivos es en realidad mínimo tomando en cuenta la cantidad de registros que se procesan.

3.4 Reportes resultantes

Los reportes que se obtienen después de ejecutar los procesos descritos son los siguientes:

- Desglose de Llamadas
- Verificación de Secuencia (formato texto)

3.4.1 Reporte de Desglose de Llamada

Para el desglose de llamadas se genera un archivo en formato Excel “*.xls”. Se presenta un ejemplo del reporte resultante de un análisis utilizando el archivo “14306.tax” de la central Apodaca. Los datos que incluye el reporte son básicamente:

- Nombre de la central (proporcionado por el usuario)
- Fecha de emisión del reporte
- Número de llamadas por fecha/hora
- Totales de llamadas.

Central: APODACA

Fecha: 27/Ago/03

Archivo	Fecha	Hora																								Total	
		0:00	1:00	2:00	3:00	4:00	5:00	6:00	7:00	8:00	9:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00		
14306	25/08/03	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	5	4	6	5	1	5	9	22	58
	26/08/03	19	18	10	7	8	12	1318	6129	19482	45862	57706	26600	0	0	0	0	0	0	0	0	0	0	0	0	0	157171

3.4.2 Reporte de Verificación de Secuencia

Para la verificación de secuencia se genera un archivo de texto con el nombre del archivo de tarificación y la extensión “.log”. Los datos que contiene el archivo son:

- Nombre de la central
- Hora de inicio de proceso
- Número de registro²⁴ inicial
- Número de registro donde se presentan saltos en secuencia (en caso de que los hubiere)
- Número de registro final
- Hora de fin de proceso

El siguiente ejemplo es el resultado de un archivo que no tuvo saltos en la secuencia de los registros.

```
Central: APODACA  
Inicio: 17:45:38  
1er número en Secuencia: 9913C5  
  
Último número en Secuencia: 9B79F2  
Fin: 17:45:39  
Secuencia de registros correcta.
```

²⁴ Los números de registro se presentan en formato Hexadecimal debido a que así fué requerido por el cliente

Se presenta también un ejemplo de un archivo que presenta saltos en la numeración.

```
Central: APODACA
Inicio: 18:17:03
1er número en Secuencia: 420464

Error en secuencia: 430690
                    -->4306AA
Error en secuencia: 43069C
                    -->430611

Último número en Secuencia: 443F1D
Fin: 18:17:04
```

3.5 Implementación

La implementación es el proceso de asegurarse que el sistema sea operacional y lograr que los usuarios tomen el control de la operación para su uso como lo señala Kendall (1997, p.821).

Para la fase implementación del sistema Tax Converter se realizaron los siguientes procesos:

1. Proceso de prueba
2. Instalación de la interfase gráfica en las PCs de los usuarios
3. Capacitación de usuarios

3.5.1 Proceso de Prueba

Todos los programas de aplicación recientemente escritos o modificados de un sistema deben ser probados en forma exhaustiva.

Para el sistema se realizaron pruebas de escritorio tomando como ejemplo archivos recuperados de diferentes centrales telefónicas. Para esto, se utilizaban bloques de los archivos y se comparaban los resultados obtenidos del sistema, con los calculados de manera manual.

El proceso resultó ser tedioso, debido a que, para verificar el buen funcionamiento del sistema se tuvieron que realizar un gran número de

ejercicios. Pero el asegurar la calidad del sistema es de alta prioridad aunque el tiempo de entrega del sistema se incremente.

3.5.2 Instalación de la interfase gráfica en las PCs de los usuarios

Como se menciona en el punto 3.1.1. el sistema debe ser instalado en todas y cada una de las computadoras que se encuentran en el CNS, las cuales comparten las mismas características tanto de hardware como de software.

Para poder instalar la aplicación en dichas máquinas es necesario crear un programa de instalación. Visual Studio 6.0 incluye una aplicación llamada "Package and Deployment Wizard" (Figura 18), la cual está diseñada para generar esta clase de programas de una manera fácil y rápida.



Figura 18. Package and Deployment Wizard

Esta aplicación creará un programa de instalación en una ruta específica con un archivo ejecutable "setup.exe" que nos servirá para instalar el programa en cada una de las computadoras. El programa de instalación generado, puede ser ejecutado desde un directorio de la red LAN²⁵ a la que están conectadas todas las computadoras de la compañía, también puede copiarse a un CD y ser ejecutado desde el mismo.

Al ejecutar el "setup.exe" el sistema se instalará en la ruta que el usuario defina. Puede crearse un "Shortcut" en el escritorio de Windows para facilitar su ejecución.

Una vez terminada la instalación en cada una de las computadoras del CNS, el sistema está listo para ser operado por el personal previamente capacitado.

²⁵ LAN. Local Area Network existente dentro de la compañía.

3.5.3 Capacitación de usuarios

Dentro del CNS existe una persona que se encarga de dar soporte técnico en cuanto a la instalación y mantenimiento del software utilizado en el área, por lo que antes que nadie, es quien debe ser capacitado para la operación de este nuevo sistema, además esta persona será capaz de dar soporte y capacitar a los demás usuarios.

La capacitación se impartió a 18 usuarios del CNS, entregando a cada uno la documentación referente al 'Manual de Usuario'.

Capítulo IV. Discusión de resultados

*El miedo al fracaso no debe ser nunca la razón
para no intentar algo.*

Frederick Smith

Capítulo IV. Discusión de Resultados

Después de haber implementado el sistema ya puede ser operado por los usuarios capacitados. La siguiente tarea es la revisión de los reportes generados por el sistema así como su correcta interpretación.

4.1 Interpretación de reportes obtenidos

Antes de dar una explicación de los resultados es importante responder a la interrogante ¿En qué caso se solicita un análisis?. Existe una herramienta que realiza el cálculo de llamadas totales para cada día tomando los datos directamente de la central, después compara este total con números predefinidos que representan un promedio basado en la estadística de manera que, el total de llamadas registradas en la central en un día específico, no debe estar por debajo de los umbrales predefinidos, en tal caso se solicitará el análisis detallado de los archivos de tarificación. Recordemos que, como se menciona en el punto 2.1, los archivos se sobrescriben al menos cada 7 días, por lo que la solicitud del análisis deberá hacerse de manera inmediata con el fin de evitar la pérdida del archivo a procesar.

Para dar una explicación detallada de los resultados obtenidos se eligieron algunos ejemplos de centrales de las cuales se tenían sospechas de pérdida de registros de llamadas.

ESTA TESIS NO SALE
DE LA BIBLIOTECA

4.1.1 Ejemplo 1: Central Mayorazgo

Como primer ejemplo tomaremos el caso de la central "Mayorazgo" sobre la cual se reportaban pérdidas el día 23 de agosto del 2003.

El primer paso es la recuperación del archivo utilizando el protocolo FTP. Es recomendable crear previamente un fólдер con el nombre de la central que se procesa, abrir una ventana *command* y posicionarse en el fólдер donde se guardará el archivo recuperado. Posteriormente abrir una sesión FTP con la dirección IP correspondiente a la central²⁶ y dar los comandos descritos a continuación²⁷:

Se abre la conexión FTP:

```
C:\Taxation\Mayorazgo\ftp
ftp>open 13.20.35.227
Connected to 13.138.55.225
220 ftp server (Monday, Ago 23 12:12:34 EST 2003) ready.
user: oper1
331 password required for oper1.
Password: *****
user oper1 logged in
ftp>_
```

Se verifican las fechas de los archivos existentes, para seleccionar el archivo adecuado. Este paso es muy importante ya que por el tamaño de los

²⁶ Las direcciones IP de las centrales se obtienen consultando un archivo que se encuentra disponible en la red interna de la compañía y todos los usuarios tienen acceso read-only.

²⁷ Los caracteres marcados con letras en negro representan los comandos tecleados por el usuario.

En promedio la transferencia del archivo dura entre 3 y 5 horas²⁸. Ya con el archivo transferido, procedemos a analizarlo ejecutando la aplicación "Tax_Converter". El reporte obtenido se muestra en la siguiente página (Figura 19).

Como podemos notar, el archivo procesado contiene registros de llamadas a partir de las 10:00 horas del día 22 de Agosto hasta las 13:00 horas del día 23 de Agosto. Observando en el reporte los datos que corresponden al 22 de Agosto, se muestran muy pocas llamadas entre las 13:00 y las 19:00 horas, así mismo, el día 23 de febrero a las 13:00 horas se registran cero llamadas; por estadística se sabe que en estos horarios la central procesa un número importante de llamadas ya que son consideradas "horas pico", por lo tanto, podemos concluir que efectivamente existe una pérdida de registros en la central Mayorazgo.

El reporte de Desglose de Llamadas nos proporciona información relevante acerca de la hora en que ocurrió la pérdida de registros de llamadas, pero no es posible conocer el número exacto de llamadas no registradas a través de este reporte. El proceso de verificar si se perdió la secuencia en el número de registro consecutivo nos podría dar un dato más exacto acerca de cuantas llamadas no se registraron en el archivo de tarificación, es de suma importancia aclarar que el hecho de que se hayan reportado pérdidas de registros en un rango de tiempo determinado, no necesariamente implica una pérdida en el consecutivo de los registros, el que esto ocurra depende del origen del problema que tuvo la central. De cualquier manera la verificación de secuencia debe ser aplicada a todos los archivos que se hayan solicitado procesar.

²⁸ Esta variación depende de la velocidad de conexión a la red y del tamaño del archivo.

MEMORIA DEL DESEMPEÑO

Central: MAYORAZGO																										
Fecha: 26/Ago/03																										
Archivo	Fecha	Hora																					Total			
		0:00	1:00	2:00	3:00	4:00	5:00	6:00	7:00	8:00	9:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00	
14300	22/08/03	0	0	0	0	0	0	0	1	0	0	1603	15107	13071	1	2	2	6	32	88	3126	15413	14548	11722	4539	89261
	23/08/03	1433	689	487	511	357	366	723	2561	6677	11764	3046	15107	1499	8876	0	0	0	0	0	0	0	0	0	38988	
		1433	689	487	511	357	366	723	2561	6677	11764	14649	15107	14570	8877	2	2	6	32	88	3128	15413	14548	11722	4539	128249

Figura 19. Reporte Central Mayorazgo

A continuación se muestra el resultado de la verificación de secuencia de la central Mayorazgo:

Central: MAYORAZGO
Inicio: 11:20:55
1er número en Secuencia: 552088
Último número en Secuencia: 571581
Fin: 11:20:56
Secuencia de registros correcta.

En este caso no se reportó pérdida alguna en el número consecutivo de registro, el primer y último número se presentan en formato hexadecimal por que así lo requirió el área de Soporte Técnico de la compañía. Podemos verificar el resultado de la siguiente manera:

Convertimos a decimal los números inicial y final:

$$\text{DEC}(552088) = 5578888$$

$$\text{DEC}(571581) = 5707137$$

Calculando la diferencia entre ambos tenemos el total de llamadas:

$$5707137 - 5578888 = 128249$$

Comparando el resultado con el total reportado en el Desglose de Llamadas notamos que son iguales. Lo cual confirma que ocurrió una pérdida de registros de llamadas pero el consecutivo no se vió afectado²⁹.

4.1.2 Ejemplo 2 Central Ocotlán

Presentando como un segundo ejemplo el reporte de resultados de la central "Ocotlán"³⁰. En la siguiente página se muestra el reporte resultante después de procesar el archivo 14345.tax (Figura 20).

Puede notarse en el reporte que, entre las 6:00 y las 11:00 horas se reportan pocas llamadas considerando que son horas de mucha actividad laboral.

²⁹ El hecho de que se vea afectado o no el número consecutivo, proporciona al área de soporte una base para investigar cuál pudo ser el origen del problema en la central.

³⁰ Omitiremos el paso de la transferencia ya que se ha descrito previamente.

Se ejecuta el proceso de verificación de secuencia con el siguiente resultado:

```
Central: OCOTLAN
Inicio: 01:24:59
1er número en Secuencia: A0E5EF

Error en secuencia: A28775
-->0
Error en secuencia: 0
-->A28789
Error en secuencia: A287A0
-->FFFB
Error en secuencia: FFFB
-->FFFB
Error en secuencia: FFFB
-->13B
Error en secuencia: 13B
-->210B
Error en secuencia: 210B
-->13B
Error en secuencia: 13B
-->FFFB
Error en secuencia: FFFB
-->FFFB
Error en secuencia: FFFB
-->13B
Error en secuencia: 13B
-->FFFB
Error en secuencia: FFFB
-->FFFB
Error en secuencia: FFFB
-->13B
Error en secuencia: 13B
-->13B
Error en secuencia: 13B
-->210B
Error en secuencia: 210B
-->0
Error en secuencia: 0
-->A287C6

Último número en Secuencia: A2FA56
Fin: 01:25:00
```

El reporte muestra varios saltos en el número de secuencia, en este caso no hay concordancia en los saltos, es decir, existen números que no se relacionan con la secuencia que lleva el archivo. En estos casos se dice que hubo una corrupción de datos dentro del archivo.

Si calculamos la diferencia el total de llamadas basándonos en el primer y último número de secuencia del archivo:

$$\text{DEC(A0E5EF)} = 10544623$$

$$\text{DEC(A2FA56)} = 10680918$$

Calculando la diferencia entre ambos tenemos el total de llamadas que debieron registrarse:

$$10680918 - 10544623 = 136295$$

Al comparar el total de llamadas el resultado con el total reportado en el Desglose de Llamadas tenemos que:

$$136295 > 135428$$

Tenemos una diferencia de 867 registros que no se reportaron en el Desglose, por lo tanto, se confirma que hubo una pérdida de registros de

llamadas en la central. Aunque se percibe una pérdida de 867 registros, se puede suponer que el número de registros perdidos pudo ser mucho mayor.

Los ejemplos presentados pretenden mostrar de manera genérica la aportación del sistema Tax_Converter al mecanismo de solución aplicado. Por su puesto que estos ejemplos son una mínima muestra de un conjunto de problemas que se presentan cotidianamente en la central.

4.2 Mantenimiento del sistema

Existen dos factores que implicarían una modificación al sistema "Tax Converter":

1. El sistema funciona exclusivamente con datos provenientes de la Central Telefónica S12 Release 3.2. El tamaño del CDR en esta versión es de 110 bytes. En tanto, las futuras versiones del Software no alteren ya sea el tamaño o la estructura del CDR el sistema podrá seguir siendo operado sin alteración alguna. Cuando el cambio en la versión involucre un cambio en el CDR, será necesario modificar principalmente los programas que generan el desglose y la secuencia con base a la nueva estructura del CDR.
2. En el momento en que se requiera algún cambio en el formato del reporte se deberá modificar exclusivamente la subrutina que envía los datos a la hoja de Excel.

Como se indica en el capítulo I, la solución funciona exclusivamente para centrales telefónicas S12 instaladas dentro del territorio nacional, sin duda, los problemas de tarificación pueden presentarse en centrales telefónicas S12 instaladas en otros países; para poder utilizar el sistema "Tax Converter" en dichas centrales, habría que realizar un análisis de la estructura tanto del CDF como del CDR correspondientes y en tal caso, realizar las modificaciones necesarias al programa para poder leer los datos y así generar el reporte correspondiente.

Conclusiones

*Tú eres el único que puede hacer la diferencia.
Cualquiera que sea tu sueño, ¡Ve por él!*
Magic Johnson

Conclusiones

El diseño del sistema "Tax Converter" fue una solución viable para optimizar el proceso de manipulación de archivos de tarificación, proceso que se realizaba de forma manual. Anteriormente, la entrega de resultados tomaba aproximadamente cuatro días, ya que era necesario asegurarse de que los resultados eran confiables y debían ser revisados más de una vez; con la utilización del sistema, las respuestas a las solicitudes de reportes se enviaban a más tardar día siguiente de la recepción de la solicitud.

El tiempo de generación de reportes utilizando el sistema se redujo considerablemente, ya que cada reporte tarda en generarse menos de un minuto. El único proceso para el que no se pudo reducir el tiempo fue el de la recuperación de los archivos de tarificación, ya que el cliente sólo autorizó el acceso a su red a través de una conexión por módem.

Una vez entregados los reportes al área correspondiente, es posible tomar una decisión con base en los resultados y lo más importante, investigar cuál fue el evento que pudo provocar una pérdida de datos en la central.

En el Centro Nacional de Servicio se recibe un número importante de solicitudes no sólo relativas a la tarificación, estas solicitudes se van asignando conforme a la hora en que llegaron y a su prioridad, por lo general, el personal a cargo de atenderlas no es suficiente en número para dar respuestas inmediatas a cada una de ellas, lo que genera a la compañía costos monetarios (horas-hombre) y la imagen ante el cliente se deteriora. Por razones de interés monetario, los problemas detectados en procesos de

tarificación son clasificados con prioridad "alta" y requieren atención y respuestas rápidas.

El invertir tiempo y costos tanto en el estudio de la información, como en el desarrollo del sistema, resultó provechoso ya que implicó una mejoría en los tiempos y procedimientos de respuesta a las solicitudes del cliente.

Otro de los beneficios que derivaron de este desarrollo fue que la confianza en la capacidad de los sistemas para manipular los datos de las centrales se incrementó y el sistema "Tax Converter" fue sólo uno de los muchos desarrollos que se implementaron para hacer más eficientes muchos procesos importantes que implican la manipulación de los datos de las centrales telefónicas S12-40.

Anexos

Anexo A. Glosario

Abonado. Usuario suscrito a una compañía que proporciona servicios de telefonía.

CDF. Call Data File. Archivo que contiene datos de llamadas.

CDR. Call Data Record. Registro de datos de una llamada.

Central Telefónica / Central. Equipo que cuenta con la infraestructura y tecnología necesarias para concentrar un grupo de abonados interconectados entre los cuales pueden establecer una comunicación ya sea de voz o datos.

CNS. Centro Nacional de Servicio

EXCHANGE. Central Telefónica.

GUI. Graphical User Interface.

IP Address. Internet Protocol Address

PSTN Public Switched Telephone Network

RCD. Red Corporativa de Datos.

RTPC. Red Telefónica Pública Conmutada.

Subscriber. Usuario suscrito a una compañía telefónica

Tarificación. Proceso de cobro de llamadas a los abonados.

WAN. Wide Area Network. Red de área amplia.

Anexo B. Lista de Figuras

Figura 1. Red Telefónica Pública Conmutada.

Figura 2. Interconexión entre centrales.

Figura 3. Las centrales a lo largo del país.

Figura 4. La central Telefónica.

Figura 5. Unidades de disco.

Figura 6. Red Corporativa de Datos RCD.

Figura 7. Proceso de desarrollo.

Figura 8. Recuperación remota de datos.

Figura 9. Formato del reporte.

Figura 10. Diagrama de flujo: Transferencia de archivos.

Figura 11. Diagrama de flujo: Desglose de llamadas.

Figura 12. Visual Basic Reference.

Figura 13. Diagrama de flujo: Envío de resultados a Excel.

Figura 14. Diagrama de flujo: Verificación de Secuencia.

Figura 15. Interfase gráfica de usuario.

Figura 16. Mensaje.

Figura 17. Avance de procesamiento.

Figura 18. Package and Deployment Wizard.

Figura 19. Reporte Central Mayorazgo.

Figura 20. Reporte Central Ocotlán.

Bibliografía

- Camacho Cansino, S. (1998). **Análisis de Algoritmos**. México: UNAM ENEP Acatlán.
- Cohen, A. (1985). **Estructura Lógica y Diseño de Programas**. Madrid: Paraninfo.
- Cornell, Gary. (1999). **Visual Basic 6.0: Manual de Referencia**. Madrid: McGraw-Hill.
- Gerez, Victor. (1988). **El Enfoque de Sistemas**. México: LIMUSA.
- Kendall, Kenneth y Kendall, Julie (1997). **Análisis y Diseño de Sistemas**. México: Prentice Hall.
- López, Leobardo. (1994). **Programación Estructurada**. México: Computeq.
- Microsoft Press. (1998). **Microsoft Visual Basic 6.0 programmer's guide**. Redmond WA: Microsoft Press.
- Moritsugu, S. y DTR Business Systems, Inc. (2000). **UNIX. Serie Práctica**. Madrid: Prentice Hall, Pearson Education.
- Verzello, R. J. y Reuter, J. (1983). **Procesamiento de Datos**. México: McGraw-Hill