



ORGANISMO COORDINADOR DEL SISTEMA ELÉCTRICO NACIONAL INTERCONECTADO DE LA REPÚBLICA DOMINICANA

INFORME DIAGNÓSTICO DE LA SITUACIÓN ACTUAL

Preparado por:

Hugo Alejandro De la Cadena Galarza
Consultor proyecto Data Warehouse

Santo Domingo, Septiembre 8 del 2006



TABLA DE CONTENIDOS

| | |
|--|-----------|
| 1. INTRODUCCIÓN | 3 |
| 2. OBJETIVO DEL DISEÑO DEL DATA WAREHOUSE | 3 |
| 3. OBJETIVOS DEL PRIMER INFORME..... | 4 |
| 4. PROBLEMÁTICA ACTUAL..... | 4 |
| 5. SITUACIÓN ACTUAL DE LA BASE DE DATOS DEL OC..... | 6 |
| 6. INFORMACIÓN DE OTRAS INSTITUCIONES DEL SECTOR ELÉCTRICO..... | 13 |
| 7. SOFTWARE EXISTENTE EN EL OC..... | 14 |
| 8. SEGURIDADES Y ACCESO A LA INFORMACIÓN..... | 17 |
| 9. OBSERVACIONES SOBRE EL MANEJO DE INFORMACIÓN Y LAS NECESIDADES ACTUALES..... | 18 |
| 10. BENEFICIOS DE ELABORAR UN DATA WAREHOUSE..... | 21 |
| 11. CONCLUSIONES Y RECOMENDACIONES..... | 21 |



1. Introducción

Sobre la base del Proyecto de Asistencia Técnica al Sector Energético, se ha contratado la consultoría “Diseño del Data Warehouse para el Organismo Coordinador del Sistema Eléctrico Nacional Interconectado de la República Dominicana con el fin de establecer un sistema global, centralizado, altamente organizado de almacenamiento de datos con capacidades de pre-procesamiento, que absorba e integre la información de las bases de datos que ya existen o se están elaborando en el OC, incorpore información almacenada en el OC empíricamente en hojas y libros Microsoft® Excel y, además, permita agregar otra información histórica del sector almacenada en otras organizaciones.

Siendo el OC i) El centro que capta gran cantidad de datos y que produce permanentemente mucha otra información relacionada con el SENI, y ii) la entidad a la que las organizaciones mencionadas recurren por información y lo seguirán haciendo en el futuro previsible, ésta institución necesita contar con una herramienta ágil y un sistema altamente organizado para el almacenamiento a gran escala de toda la información cuantiosa que le llega en tiempo real, para la minería y almacenamiento efectivos y eficaces de data histórica necesaria para muchas proyecciones, para el pre-procesamiento de información, y para consultas, búsquedas, análisis estadísticos, y elaboración de reportes en tiempos razonables y adecuados.

El proyecto del Data Warehouse comprende el almacenamiento de información estructurada (Bases de Datos) y no estructurada (Archivos) A esta fusión de información que va a ser contenida en un solo repositorio central se le ha dado el nombre de “Sistema Integral e Integrado de Manejo y Administración de Datos”

Este documento abarca la primera etapa del proyecto, que es: El diseño del Data Warehouse, la situación actual, la arquitectura y los términos de referencia que servirán de base para la segunda etapa del proyecto, la construcción del Data Warehouse.

2. Objetivo del Diseño del Data Warehouse

El Data Warehouse tiene por objetivo almacenar información en una base de datos histórica y centralizada. Esta información sirve para la elaboración de reportes gerenciales y toma de decisiones.

3.1 Objetivos Específicos



- Permitir que la información de la organización sea accesible desde un nivel agrupado (gerencial) hasta un nivel de detalle (operativo)
- Generar reportes ejecutivos con rápido desempeño, sin tiempos de procesamiento y sin tiempo de espera.
- Contener una sola versión de la información del Organismo Coordinador, en un solo repositorio central, sin duplicidad y de alta calidad.
- Contar con información del sector eléctrico, organizada de manera centralizada, que permita apoyar la construcción de modelos analíticos de los datos del mercado, generar proyecciones, analizar patrones del mercado, determinar patrones, visualizar tendencias, etc.

3. Objetivos del primer informe

El Diseño del Data Warehouse está dividido en 3 informes que van a ser elaborados en esta primera etapa. Este documento contiene el producto esperado No. 1: Informe Diagnóstico sobre la situación actual de la base de datos del OC, las herramientas de software existentes y los requerimientos de información de la Alta Gerencia del OC, los Agentes del Sector, y las organizaciones formativo-reguladoras del Sector Eléctrico Dominicano.

4. Problemática Actual

La mayoría de la información del Organismo Coordinador se almacena de manera no auditada, es decir, en archivos Excel, archivos de texto, bases de datos satélites que no se encuentran interconectadas. Esto ocasiona que no haya una fuente confiable de la información.

El sector no tiene una institución dedicada a elaborar estadísticas eléctricas ni proyecciones que pueda proveerlas en tiempos razonables a los propios agentes del sector, a organismos normativo-reguladores, entre otros, con fines de análisis y tomas de decisiones formativas, reguladoras, operativas o de inversión.

Las problemáticas actuales se pueden resumir en:



Formato de almacenamiento

La mayoría de los archivos que guardan información lo hacen a manera de matrices y no a manera de tablas o vectores. Para almacenar esta información en una base de datos relacional se debe transformar el formato de esta información. Esto se debe hacer, no solamente como un requisito del Diseño del Data Warehouse (OLAP), sino como un requisito para un sistema transaccional (OLTP)

Ausencia de sistema transaccional integrado

Actualmente no se cuenta con un sistema automatizado en donde se puedan hacer consultas de información sin intervención de las áreas que generan dicha información, esto hace que los departamentos que generan información sean también aquellos que deben generar los reportes sobre la data terminada. Esa labor debe ser automatizada.

Duplicidad de información

Al no contar con una herramienta centralizada existe una gran cantidad de duplicidades de información, cada usuario graba sus archivos Excel para su trabajo diario. Esto ocasiona, varias versiones de la realidad y almacenamiento innecesario de la información duplicada.

Limitaciones en crecimiento

Lo que actualmente se maneja en Excel está creciendo de tal manera que ésta herramienta está alcanzando su capacidad de almacenar la información. Actualmente la cantidad de columnas del Excel está siendo utilizada casi en su totalidad por muchos de los archivos del trabajo diario.

Información delicada

Un archivo Excel, o archivo de texto es muy sensible a errores humanos, en la mayoría de los casos los archivos permiten el ingreso de cualquier tipo de información, no hay controles en la lectura. Esto puede ocasionar errores en generación de información y se traduce en doble trabajo para estas correcciones.



Sistema de toma de decisiones

No existe un sistema de información gerencial que provea información de nivel gerencial de una manera ágil y oportuna. Cuando se hace un requerimiento de información gerencial histórica no hay una manera automática de hacerlo. **Esto es principalmente el objetivo del proyecto del Data Warehouse.**

5. Situación Actual de la base de Datos del OC

En el Organismo Coordinador se cuenta con varios paquetes de Software que permiten el trabajo de varias divisiones. El problema fundamental es que estos sistemas cubren un pequeño porcentaje del trabajo diario, el resto está automatizado en Excel.

Existen sistemas que permiten el trabajo diario, pero no son almacenadores de información histórica, sino que cumplen con labores específicas, reciben insumos, aplican una serie de algoritmos y arrojan un resultado, pero la información que generan nuevamente se almacena en Excel y en varios casos a Bases de datos de manera manual. Este es el caso de Startnet, DIGSILENT por ejemplo.

Los archivos Excel no son utilizados de una manera común y corriente sino que literalmente los mismo usuarios han hecho pequeños sistemas para la automatización de sus informaciones en esta herramienta. Esto indica que en ausencia de un sistema transaccional, los usuarios han resuelto una parte de sus necesidades automatizando los archivos Excel mediante macros.

De acuerdo a las divisiones de los departamentos, a continuación se especifica los insumos, procesos e información terminada que elabora cada departamento.



División Programación de Operaciones

Insumos de Información

Programa Semanal de Operación

Tasas de cambio

Mantenimientos de circuitos de distribución

Reportes de precios de combustibles

Demanda de distribuidores

Disponibilidad Unidades Térmicas

Disponibilidad Unidades Hidroeléctricas

Costo Variable de Producción

Actualizaciones de diagrama unifilar

Demanda de Usuarios No Regulados

Histórico de demanda del Sistema

Histórico de demanda Barra

Procesos y Software utilizado

Programación de Corto Plazo (Diario y Semanal)

Programación de Mediano Plazo

Programación de Largo Plazo

Elaboración de Programa de Mantenimiento Mayor

Software Utilizado

Energy Audit

Moperd

Starnet

Programación Diaria (Manzuela Net)

DIGSILENT

PSS/E

Información Terminada

Programa Diario de Operación

Programa Semanal de Operación

Programa de Operación de Largo Plazo

Reporte de Disponibilidad por Unidad

Lista de Mérito de despacho

Factores de Nodo

Energía No suministrada

Programa de Mantenimiento Mayor

Programa Mediano Plazo

Informe Mensual



División de Supervisión en Tiempo Real

Insumos de Información

Programa Diario Operación

Lista de Mérito

Información de Generadoras (Características)

Potencia horaria de generadoras

Monitoreo visual y auditivo

Niveles de Frecuencia mínimo y máximo

Reporte de Indisponibilidad previsto

Procesos y Software utilizado

Carga Reporte Diario de Operación

Carga Secuencia de Eventos

Reporte de Indisponibilidad por Unidad

Software Utilizado

RDO
SCADA
Spread Sheet
SIGRA

Información Terminada

RSE - Secuencia de Eventos

RDIU - Reporte de Indisponibilidad por Unidad (Real)

Informe Diario de Eventos Relevantes

Informe de Incumplimientos y desacatos

Informe de Fallas

Informe de Regulación Primaria y Secundaria de Frecuencia

Máquinas forzadas por tensión

Informes sobre desacoples e islas económicas

Indicadores de Calidad de Frecuencia

Reporte Preliminar TR

Factores de Eficiencia



División de Estudios

Insumos de Información

Programación diaria

Demanda del sistema

Oferta de generadoras

Mantenimientos de Líneas

Ingresos de nuevos equipos

Ubicaciones sobre nuevos elementos

Transacciones Económicas

Histórico sobre demanda del sistema

RDO - Reporte Diario de Operación

RSE - Secuencia de Eventos

Procesos y Software utilizado

Estudios de Viabilidad y Estabilidad

Estudios de Factibilidad

Simulaciones

Estudios a Largo Plazo

Software Utilizado

Moperd
StarNet
SIGRA
PSS/E
DIgSILENT

Información Terminada

Plantas forzadas para soporte de Tensión

Informes de Estudios a mediano Plazo

Informes de Estudios a Largo Plazo

Informes de causas sobre problemas en el sistema

Archivos de estudios varios realizados

Informes sobre la factibilidad de inclusiones de nuevos elementos en el Sistema

Estudio Corto Circuito

Base de Datos de Protecciones



División de Mediciones

Insumos de Información

Registros obtenidos de los medidores

Registros recibidos de los agentes

RDO - Reporte Diario de Operación

RSE - Reporte de Secuencia de Eventos

Factores de nodo

Informe de Visitas

Formulario de Administración

Salidas al terreno

Información suministrada por los Agentes sobre los SMC

Procesos y Software utilizado

Captura de información de Medidores

Validaciones entre mediciones enviadas y capturadas

Almacenamiento de información de visitas

Software Utilizado

Energy Audit
SADOC
ION Enterprise
Nexus
Laserfiche
Proyecto PROTEO_V (Desarrollo)

Información Terminada

Mediciones validadas

Registros de Visitas

Comparaciones entre registros obtenidos y recibidos

Informe mensual SADOC

Informe mensual SMC

Base de Datos de los Medidores



División de Transacciones

Procesos y Software utilizado

Insumos de Información

Información Terminada

Mediciones Validadas (Reales)

Informes de Operadores

RDO - Reporte Diario de Operación y RSE - Secuencia de Eventos

Contratos (Cantidad de Energía y Potencia)

Pronóstico de Demanda Máxima y Demanda de Potencia de Punta

Matriz de Medidores y Matriz de Contratos

Demanda Máxima Anual

Costos Marginales de Energía por barra

Costos Variables de Producción y Despacho

Resoluciones que fijan: posibles compensaciones de energía por costo marginal máximo, Otras compensaciones y Costo por desabastecimiento

Valores máximo de los Costos Marginales, Consumer Price Index

Resolución que fija el Incentivo de RF y Compensaciones por RF

Parámetros de la Red y de las Unidades

Estadística de Indisponibilidades para el cálculo de Potencia Firme

Factores de Nodo de Energía y de Potencia

Desacoplos físicos y económicos

Tasa de Interés y Tasa de cambio del Banco Central

Topología de la Red

Calculo de Transacciones Económicas de Energía

Calculo de Transacciones Económicas de Potencia

Cálculo de compensaciones de derecho de conexión

Cálculo de compensaciones por regulación de frecuencia

Calculo de los Costos Marginales

Cálculo de Demanda Máxima

Indisponibilidad para el Cálculo de potencia Firme

Factores de nodo de Potencia

Software Utilizado

PROTEO_E (Desarrollo)

PROTEO_V (Desarrollo)

PFT

Laserfiche

PSS/E

Factores de Nodo de Potencia

Transacciones Económicas de Energía Activa

Transacciones Económicas de Potencia de Punta Preliminar

Transacciones Económicas de Potencia de Punta Definitivas

Transacciones Económicas por Derecho de Conexión preliminar

Transacciones Económicas por Derecho de Conexión Definitiva

Compensaciones por Regulación de Frecuencia

Costos Marginales de Corto Plazo de Energía y de Potencia de Punta

Otras Compensaciones

Transacciones Económicas de Energía Reactiva

Base de Datos de Contratos



Gerencia de Administración, Finanzas y Tecnología

| Insumos de Información | Procesos | Software | Información Terminada |
|---|---|------------|--|
| Ingresos y salidas de personal | Sistema de Control de Acceso | Integra32 | Detalle de los ingresos y salidas de personal |
| Información de personal interno y candidatos, expedientes, capacitaciones | Sistema de Recursos Humanos | Eikon | Información del personal (área laboral, nivel académico, edad, tiempo en institución, reconocimientos, etc.) |
| Ingresos y movimientos contables | Sistema Financiero y Contable | DacEasy | Detalle de plan contable |
| Presupuestos | | | |
| Llamadas telefónicas salientes | Sistema de Control de Llamadas | Avotus | Detalle de las llamadas salientes (extensión, numero, duración, hora, etc.) |
| Documentos para registros | Digitalización de Documentos | Laserfiche | Documentos digitalizados en base de datos binaria |
| Mantenimiento de Activos fijos, inventario de suministros y material gastable | Activos Fijos, Material Gastable y Suministros | Excel | Archivo de control de Activos fijos, material gastable y suministros |
| Asignación de permisos de acceso a los sistemas | Control y Asignación de permisos de Acceso a Sistemas | Windows | Permisos asignados por usuario, por sistema |



6. Información de otras instituciones del sector eléctrico

Para fines de esta etapa y como se mencionan en los términos de referencia, cuando se mencionen las “otras instituciones del sector eléctrico dominicano” se refiere a:

- Hidroeléctrica CDEEE
- Comisión Nacional de Energía (CNE)
- Superintendencia de Electricidad (SIE)
- Corporación Dominicana de Empresas Eléctricas Estatales (CDEEE)
- Sistema de Transmisión CDEEE

Además de las informaciones que las otras instituciones del sector entregan al Organismo Coordinador y que se especifican en las leyes y reglamentos del sector eléctrico; este primer informe especifica otras adicionales que el OC requiere –no necesariamente- para realizar un mejor análisis de su desempeño.

Estas informaciones deben ser revisadas por las otras instituciones del Sector Eléctrico para validar si se encuentran almacenadas en un formato adecuado para realizar una lectura automática y además, estas instituciones deben entregar un visto bueno sobre si están de acuerdo en compartirlas. Las informaciones divididas por institución son las siguientes:

- **Hidroeléctrica CDEEE**
 - Características de los Embalses
 - Altura disponible
 - Altura libre
 - Cota Mínima
 - Cota Máxima
 - Volumen Útil
 - Cota horaria
 - Caudal promedio diario
 - Material de la presa
 - Características de la ubicación
 - Sector
 - Climatología
 - Tipo de región
 - Cadenas Hidráulicas
 - Cuotas de uso diario (energía, riego)

- **Comisión Nacional de Energía (CNE)**



- Planes indicativos de expansión de la generación
- Proyectos de expansión de las empresas del sector

- **Superintendencia de Electricidad (SIE)**
 - Reglamentos
 - Leyes
 - Resoluciones
 - Sanciones
 - Proyectos y solicitudes de nuevas instalaciones
 - Valor agregado de distribución

- **Corporación Dominicana de Empresas Eléctricas Estatales (CDEEE)**
 - Precios de la energía EDE's (diario)
 - Precios de la energía a consumidor final (diario)
 - Precios de combustibles por tipo (diario)
 - Aportes y subsidio del gobierno
 - Indicadores macroeconómicos
 - Tipo de Cambio
 - IPC
 - Proyectos de expansión
 - Energía Facturada
 - Energía Cobrada
 - Cantidad de abonados finales
 - Consumo de energía por Sector

- **Sistema de Transmisión CDEEE**
 - Proyectos de expansión
 - Diagrama Unifilar
 - Estadísticas de fallas (Maniobras de equipos)
 - Proyectos Ejecutados

7. Software existente en el OC

Para el trabajo diario, el OC cuenta con una serie de paquetes informáticos que se describen a continuación .

SCADA

Se utiliza para el monitoreo y supervisión en tiempo de real de la actividad eléctrica del SENI. Estos son equipos y programas, forman parte integral del



sistema SCADA del SENI, el cual permite interconectar todos los componentes que integran la red de electricidad (Plantas Termoeléctricas, Hidroeléctricas, Transformadores, Barras, etc.) De manera que su estado pueda ser monitoreado facilitando así las labores de prevención, control y seguridad para el buen funcionamiento y operación del SENI en su conjunto, en coordinación con el Centro de Control de Energía (CCE).

En Organismo Coordinador posee licencias de acceso de lectura al sistema SCADA (que se encuentra en funcionamiento en el CCE) que permiten el monitoreo de la actividad del SENI.

Este sistema almacena información importante en una base de datos que puede ser explotada por el Data Warehouse.

SADOC

Sistema de Adquisición de Datos del OC, este sistema interroga y almacena los medidores que conforman los Sistemas de Medición Comercial de los Agentes del MEM, cuyos datos son utilizados posteriormente en las transacciones económicas.

De la misma manera este sistema acumula información valiosa para la carga de un Data Warehouse.

PFT

El programa PFT, ha sido implementado por la Gerencia Comercial del Organismo Coordinador y calcula la Potencia Firme de las Centrales Termoeléctricas que conforman el Sistema Eléctrico Interconectado. Corresponde a una adaptación en lenguaje C de la versión del CAPOF (Cálculo de Potencia Firme) desarrollado originalmente en FORTRAN.

Durante el periodo de entrevistas, no se habló de este sistema, aparentemente es un sistema que ayuda en el trabajo diario pero no genera información terminada, es decir, no genera información que se almacene en archivos o bases de datos.

PSS/E

Este programa es utilizado por la División de Estudios de la Gerencia de Operaciones y por la División de Transacciones de la Gerencia Comercial para la determinación de flujos de potencia.



Este software sirve para estudios, análisis y simulaciones pero no genera información en tablas, tiene una función distinta, no es un sistema transaccional que almacena información.

DIGSILENT

Es un programa equivalente al PSSE pero soportado en un ambiente más amistoso. El DIGSILENT es utilizado por la División de Estudios de la Gerencia de Operaciones y está destinado al análisis de fallas y de estudios dinámicos.

Este software sirve para estudios, análisis y simulaciones pero no genera información en tablas, tiene una función distinta, no es un sistema transaccional que almacena información.

STARTNET

Esta herramienta fue adquirida por el OC para resolver la Programación de Corto Plazo del SENI.

El resultado de las programaciones se inserta en una base de datos de manera manual.

INTEGRA32

Esta herramienta sirve para registrar las entradas y salidas del personal del OC que se realizan mediante un carné de ingreso que poseen los empleados.

Este sistema almacena información importante en una base de datos Microsoft SQL que puede ser explotada por el Data Warehouse mediante un Data Mart

EIKON

Esta herramienta sirve para la administración de Recursos Humanos donde se almacenan informaciones sobre personal interno del OC y candidatos externos.

De la misma manera este sistema acumula información valiosa para la carga de un Data Warehouse mediante un Data Mart de RRHH

DACEASY

Esta herramienta sirve para el manejo de la contabilidad del Organismo Coordinador donde se registran movimientos contables, ingresos, egresos, etc.



De la misma manera este sistema acumula información valiosa para la carga de un Data Warehouse mediante un Data Mart Contable

AVOTUS

Esta herramienta almacena todas las informaciones sobre llamadas que se hacen en el Organismo Coordinador. Actualmente está configurado para almacenar las llamadas salientes

De la misma manera este sistema acumula información valiosa para la carga de un Data Warehouse mediante un Data Mart de Control de llamadas

LASERFICHE

Esta herramienta se utiliza para la digitalización de documentos.

Este software tiene una función específica sobre la digitalización de documentos, la forma de almacenamiento la hace mediante un sistema propietario. No aplica para el almacenamiento en el proyecto de Data Warehouse

SISTEMAS OPERATIVOS (WINDOWS)

Los sistemas Operativos que se utilizan en el Organismo Coordinador son totalmente Microsoft Windows en sus diferentes versiones de escritorio o servidor. Dentro de los sistemas operativos servidores se configuran las seguridades de acceso, lectura, escritura para acceso de usuarios.

La seguridad de aplicaciones y de sistemas Operativos es propia de cada servidor. Esta información no es explotable y no se puede almacenar en el proyecto del Data Warehouse

8. Seguridades y Acceso a la información

La información está almacenada en servidores de archivos para la lectura de los usuarios mediante carpetas compartidas, pero al manejar el mismo archivo por varios usuarios pueden existir problemas de concurrencia. Por este motivo es necesario que las informaciones para consulta se encuentren en base de datos, donde se permita lectura concurrente.



De la misma manera, la creación de una Data Warehouse y su explotación mediante cubos en Excel hacen que la información se encuentre centralizada donde se puedan aplicar mejores técnicas de seguridades. No solamente en el ámbito de permisos de lectura o no-lectura, sino a niveles más detallados, dependiendo del perfil de usuario.

9. Observaciones sobre el manejo de información y las necesidades actuales

A ctualmente la información que se maneja en Excel, tiene varias debilidades al ser una herramienta de manejo de información y no de almacenamiento histórico (como una Base de Datos), se pueden observar las siguientes debilidades:

- a) **La data se almacena en forma matricial, es decir tienen valores en las filas, valores en las columnas y datos en las celdas.**

Este tipo de almacenamiento no es apto para un sistema transaccional y tampoco para un sistema de Data Warehouse. Primeramente se tiene que verticalizar esa información, por ejemplo, esta información en forma de matriz:

| Medidor/Hora | Hora 1 | Hora 2 | Hora 3 | Hora 4 | Hora 5 |
|--------------|--------|--------|--------|--------|--------|
| Medidor 1 | 34 | 36 | 38 | 40 | 39 |
| Medidor 2 | 35 | 33 | 32 | 33 | 34 |
| Medidor 3 | 67 | 65 | 68 | 70 | 66 |

Debe ser transformada a forma de vector:

| Medidor | Hora | Valor |
|-----------|--------|-------|
| Medidor 1 | Hora 1 | 34 |
| Medidor 2 | Hora 1 | 35 |
| Medidor 3 | Hora 1 | 67 |
| Medidor 1 | Hora 2 | 36 |
| Medidor 2 | Hora 2 | 33 |
| Medidor 3 | Hora 2 | 65 |
| Medidor 1 | Hora 3 | 38 |
| Medidor 2 | Hora 3 | 32 |
| Medidor 3 | Hora 3 | 68 |
| Medidor 1 | Hora 4 | 40 |
| Medidor 2 | Hora 4 | 33 |
| Medidor 3 | Hora 4 | 70 |



| Medidor | Hora | Valor |
|-----------|--------|-------|
| Medidor 1 | Hora 5 | 39 |
| Medidor 2 | Hora 5 | 34 |
| Medidor 3 | Hora 5 | 66 |

b) Existe información que no está almacenada y que se identifica con el manejo de colores.

Toda información debe almacenarse en la base de datos, no se debe dejar nada abierto para interpretación mediante leyendas en los archivos.

Por ejemplo:

| | | | | | |
|------------------|-------|-------|-------|-------|-------|
| AES ANDRES | 265.3 | 269.5 | 274.6 | 274.3 | 273.1 |
| LA VEGA | 36.8 | 36.8 | 36.8 | 36.8 | 36.8 |
| SMITH VAP | 75.2 | 75.2 | 75.2 | 48.1 | 75.2 |
| SULTANA DEL ESTE | 127.3 | 127.8 | 141.8 | 124.7 | 128.4 |
| PALAMARA | 47.1 | 50.2 | 48.1 | 47.1 | 47.1 |
| ESTRELLA DEL MAR | 59.3 | 59.9 | | 59.3 | 59.8 |
| METALDOM | 27.0 | 28.9 | 12.0 | 12.0 | 27.0 |
| CEPP 2 | 31.5 | 31.5 | 31.5 | 31.5 | 31.5 |
| MONTE RIO | 72.8 | 72.8 | | | 72.8 |

En este ejemplo existe mucha información que no se va a almacenar en una base de datos. Esta información es interpretada por quien la está utilizando. La base de datos tiene que almacenar estas leyendas.

c) No existe una información jerárquica sobre las distribuciones geográficas.

Un sistema de toma de decisiones se caracteriza también por contener información que va desde lo general hacia lo particular. En las bases de datos del OC, una de las variables que puede tener varias jerarquías es la variable "Generadora".

Actualmente la generadora tiene 2 niveles:

- Empresa
 - Generadora



Por ejemplo,

| EGE-HAINA |
|------------------|
| HAINA 1 |
| HAINA 2 |
| BARAHONA CARBON |
| HAINA 4 |
| SAN PEDRO VAPOR |
| PUERTO PLATA 1 |
| PUERTO PLATA 2 |
| HAINA TG |
| BARAHONA TG |
| SAN PEDRO TG |
| SULTANA DEL ESTE |

Para una información gerencial, sería recomendable que se pueda analizar esta variable "Generadora" también desde un punto de vista de distribución geográfica, por ejemplo:

Zona: Sur

Provincia: San Cristóbal

Sector: Haina

Empresa: EGE-HAINA

Generadora: Generadora x

Se debe analizar este tipo de jerarquía en esta y otras variables para llegar a generar una información gerencial adecuada.

d) Tipos de informaciones existentes.

Existen 2 tipos de información que necesitan ser almacenadas en el SIIMAD (Sistema Integral e Integrado de Manejo y Administración de Datos) que varían de acuerdo a su formato.

Existen informaciones:



- **Formato Matriz o Vector**

Estos tipos de informaciones son aquellas que se almacenan en hojas de cálculo como Excel o directamente en base de datos. Son informaciones que poseen características cualitativas, cuantitativas.

El almacenamiento para el SIIMAD de estas informaciones debe ser en base de datos relacionales para ser explotadas mediante el Data Warehouse.

- **Formato Archivo**

Estos tipos de informaciones son aquellas que se manejan en procesadores de texto, corresponden principalmente a Informes, minutas de visitas, archivos propios de otras herramientas como los que almacenan la topología de la red, etc.

El almacenamiento para el SIIMAD de estas informaciones debe ser mediante archivos. Estos deben ser manejados mediante un sistema de colaboración de archivos que tenga control de versiones. La división de Estudios es la que genera y mantiene este tipo de archivos.

10. Beneficios de elaborar un Data Warehouse

Primera mente los beneficios deben ser analizados desde el interior del Organismo Coordinador. El almacenamiento central de una base de datos para consultas analíticas y tomas de decisiones aporta muchos beneficios, como son:

- Soporte en la toma de decisiones
- Apoyo en la realización de estudios y proyecciones
- Contar con un almacén de datos históricos y homogéneos del SENI
- Acceso inmediato y en línea para los Agentes (24 horas x 7 días a la semana) a través de la Web.

11. Conclusiones y Recomendaciones

El manejo de información en archivos Excel y otras fuentes es bastante delicado. Actualmente existen varios proyectos dentro del OC que poco a poco van a ir quitando a Excel su protagonismo, mientras tanto se debe



hacer un doble proceso: generar la información utilizando Excel y la información terminada debe almacenarse en bases de datos. A medida que se automaticen los sistemas la escritura de información debe ser directamente a las bases de datos relacionales.

El Data Warehouse no va a servir para el almacenamiento transaccional, para este fin son estos sistemas OLTP. El Data Warehouse atiende a requerimientos de usuarios de nivel gerencial y se alimenta de los sistemas transaccionales y sus bases de datos.

Es recomendable que toda la información se almacene en bases de datos, para que luego el Data Warehouse se alimente de dichas tablas. No es recomendable que el Data Warehouse se alimente directamente de archivos Excel debido a que no contienen los formatos necesarios y son muy delicados para su automatización y esa labor es de un sistema transaccional.