

Consideraciones sobre Data Management

Por Jorge Albano

Dentro de la industria de los hidrocarburos no existe actividad donde la informática no se utilice de manera activa. La aplicación generalizada de las diferentes técnicas en exploración y producción dio lugar a una continua generación de información que por su densidad y diversidad alcanza altos niveles de complejidad. Esto hizo necesario desarrollar sofisticados sistemas especializados que facilitan el gerenciamiento de dicha información.

La permanente búsqueda de nuevas reservas constituye un hito condicionante de la actividad de exploración y producción de hidrocarburos, cuyo objetivo es reponer lo producido e incrementar el nivel de reservas explotables. Para ese fin se recurre a la aplicación generalizada de técnicas geofísicas, geológicas, petrofísicas, de ingeniería, etc.

Estas técnicas dan lugar a una continua generación de información especializada. Su densidad y diversidad alcanza tal nivel de complejidad que para su acceso y utilización en forma óptima es necesario recurrir a sistemas altamente especializados que faciliten el gerenciamiento de esta información, sistemas conocidos con el nombre genérico de *Data Management*.

Para comprender sus características y potencial implementación es necesario tener en cuenta las siguientes consideraciones:

- El exponencial desarrollo en tecnologías de *hardware* y *software* de los últimos 10 años ha logrado poner al alcance de profesionales en geociencias herramientas de interpretación y modelado que permiten aproximarse con más precisión a la modelización de una cuenca productiva o un reser-

vorio individual. La gran capacidad y velocidad de procesamiento de estas herramientas facilitan la utilización intensiva de un amplio espectro de variables características.

- Esta misma evolución de la industria de la computación ha permitido desarrollos de tecnologías en bases de datos que proporcionan el acceso a grandes volúmenes de información en forma óptima y simple. Estas tecnologías son de uso intensivo en los diferentes ámbitos de la industria de los hidrocarburos: exploración, producción, administración, economía, instalaciones, etc.
- Dentro de la problemática de las operaciones de campo el surgimiento de nuevos controladores conectados a supercomputadoras inmersas en un sistema interno de redes (Intranet), permite el control remoto de gran parte de los equipos de un campo productivo así como la transferencia de todo tipo de información generada en el campo.
- No existe actividad dentro del Petróleo donde la informática no se utilice de manera activa, ya sea dentro de las áreas técnicas como en los niveles más altos de la gestión empresarial. En todas estas actividades el factor común es el uso intensivo de infor-

mación generada en diferentes épocas, con diferentes técnicas, en diferentes soportes físicos, y en múltiples formatos para el caso de medios digitales.

- Con el objeto de poner a disposición de estas nuevas aplicaciones toda la información específica requerida, así como para asegurar tiempos de búsqueda, recolección, y captura equivalentes a los tiempos de procesamiento es que se han desarrollado las técnicas de *Data Management (DM)* permitiendo la utilización en forma óptima de diferentes aplicaciones orientadas hacia la geofísica, geología, ingeniería, gestión, administración, etc.
- La implantación exitosa de un sistema de DM exige una detallada caracterización del escenario soporte de la información, y una clara definición de los requerimientos que este sistema debe satisfacer. Sobre la base de estas dos condiciones es factible llevar a cabo esta implantación recurriendo a desarrollos integrales existentes en el mercado, o bien procediendo a un desarrollo local.

Escenario característico

Al establecer las condiciones de partida en la definición de un sistema de DM es necesario tener en cuenta, para su posterior tratamiento, el tipo de medios donde residen los datos: digitales o bien físicos (papel, acetatos, microfichas, etc.). El nivel de complejidad y los costos asociados son diferentes para ambos casos.

Se puede considerar que desde fines



Jorge Albano

de la década del '70 comienza el ingreso masivo de información en formato digital con la introducción de unidades computarizadas en los equipos de perfilaje, y el advenimiento de las nuevas técnicas de registración y procesamiento sísmico.

En el tratamiento de la información es necesario distinguir la original de campo (*Raw Data*), de aquella producto de un procesamiento y registrada no necesariamente en formato digital. Para la información digital de campo encontramos una baja dispersión de formatos ya que en lo referente a registración sísmica existen no más de cuatro formatos de uso realmente masivo, SEG A, SEG B, SEG D, SEG Y, de los cuales el último se ha impuesto ya como el estándar general dentro de lo que son secciones sísmicas procesadas.

En el aspecto de la interpretación, el escenario está dominado por un número reducido de grandes sistemas que, en general, abarcan los diferentes flujos de procesos que conducen a la determinación de un modelo de reservorio ya sea desde un punto de vista estructural como estratigráfico (determinación de horizontes, mapeos, cortes, relación tiempo-profundidad, tratamiento de ondícula, inversión de traza, etc.). Los productos finales de estos grandes sistemas son generados en formatos finales incompatibles entre sí.

En el área de perfilaje la situación es similar, habiéndose adoptado sólo los formatos BIT, LAS, LIS, DLIS. Si bien las diferentes aplicaciones utilizadas en procesamientos de registros generan datos en formatos internos propios, se lo-

gra intercambiar información mediante la utilización del formato ASCII.

Los registros de perfilaje previos a la incorporación de la computación en las unidades de campo se encuentran generalmente en papel o en acetato. Ha habido por parte de las principales empresas productoras del país un apreciable esfuerzo en la digitalización de estos perfiles con el doble objetivo de asegurar su perdurabilidad y a su vez permitir su procesamiento con las mismas herramientas que se usan para los perfiles digitales modernos.

En lo concerniente a la digitalización de viejas secciones sísmicas en papel, se ha procedido de forma análoga, digitalizándose traza por traza y obteniendo como producto final un tape en formato SEG Y similar a las secciones obtenidas de registros digitales de campo.

Una valiosísima información ya sea desde un punto de vista de la ingeniería de reservorios como de la producción reside en los legajos de pozo, donde tanto los datos numéricos, información de todo tipo, partes varios y demás documentos operativos están en textos escritos generalmente a mano. En este aspecto de la información de pozos también ha habido un gran esfuerzo por parte de las empresas productoras del país en migrar esta información al formato digital, proceso cuyo principal inconveniente es la necesidad de contar con un analista de la información con sólida experiencia en la interpretación de los datos de pozo.

Dentro de estos legajos encontramos ensayos de pozo: buildup, isocronales, falloff, análisis PVT, pruebas de

velocidad, etc. Al procederse a su digitalización es necesario incluir no sólo su interpretación final, sino también los datos registrados en el propio ensayo.

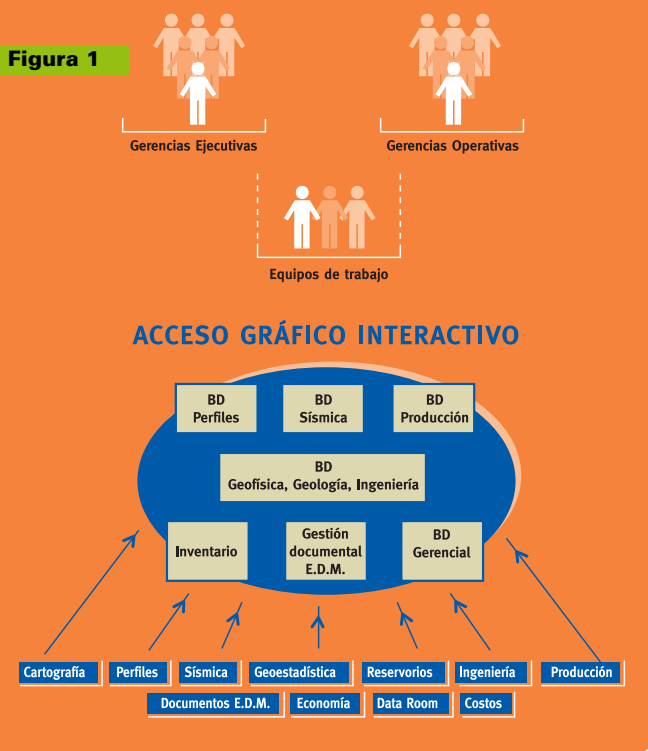
Una fracción apreciable de datos petrofísicos, ensayos de formaciones y caracterización de fluidos se encuentra volcada en forma manuscrita sobre los mismos registros de perfilaje en papel. Dado lo valioso de esta información es necesario considerar la transferencia de estos datos a medios digitales a los efectos de preservarla y a su vez incorporarla a las bases de datos respectivas.

Los datos generados a nivel de interpretación se encuentran en general dentro de documentos técnicos como son informes, reportes, memorandos, etc. Aquellos generados en forma digital poseen una multiplicidad de formatos digitales tanto para la parte de texto como para gráficos. Dependiendo de la antigüedad de la operación, muchos de estos documentos residen sólo en papel, lo que ha llevado a varias empresas a iniciar un proceso de escaneo masivo de estos documentos junto con los legajos de pozo.

Un caso análogo sucede con las secciones sísmicas en papel, donde además de la información manuscrita se encuentran impresas las tablas de velocidades de *stack*. Estas tablas deben de preservarse cualquiera sea el proceso de digitalización que se implemente sobre la respectiva sección sísmica.

La mayor densidad de información se encuentra en forma de mapas y secciones sísmicas, sobre las cuales se han volcado todo tipo de datos e interpretaciones. Los formatos son fuertemente dependientes de las aplicaciones que

Figura 1



ción de este tiempo a valores mínimos constituye uno de los requerimientos esenciales para estos sistemas.

Es necesario facilitar el acceso a la información, ya sea ésta tanto gráfica como numérica, desde el mismo puesto de trabajo donde se está implementado el flujo de trabajo inherente a un proyecto en particular. El profesional de geociencias debe poder acceder desde cualquier estación remota, en forma interactiva y al toque del dedo (*finger tip*) a cualquier tipo de datos geológicos,

cuentra totalmente generalizado el uso de coordenadas Gauss-Krueger, no se ha conservado el punto de referencia al cual están referidas (Inchauspe, Choss Malal, etc.) obligando a un trabajo de corrección adicional. El sistema de DM debe soportar el manejo de todo tipo de proyecciones cartográficas y sus respectivas transformaciones, condición necesaria para asegurar el acceso a todo tipo de proyectos independientemente de sus coordenadas originales, y además generar resultados en sistemas variados de acuerdo a las exigencias.

La utilización de una ventana de acceso (*front-end*) común para acceder a los diferentes tipos de datos aquí mencionados simplificará al geocientista la operatividad del sistema de DM, pudiéndose dedicar casi exclusivamente al propio análisis físico del proyecto en estudio y no a la recopilación de información.

Una vez finalizado todo proyecto técnico el sistema debe incluir el proceso que asegure el almacenamiento de sus resultados en las bases de datos propias y con acceso a éstos en igual forma a los datos originales.

Teniendo en cuenta la falta de un criterio racional en la nominación de los pozos detectado en prácticamente todos los yacimientos del país, es necesario establecer un criterio único de nominación y definir un proceso de creación de pozos dentro del DM. Algo similar debe establecerse para las líneas sísmicas.

Se deberá establecer un sistema integral de validación de la información que comprenderá dos aspectos bien definidos: un aspecto automático de condicionamiento de los datos mediante controles informáticos; y un aspecto necesario de validación en forma manual que exige la participación activa de los propios generadores de la información para el control de calidad de los datos. Para ello es necesario la definición de un esquema de dueños de datos.

Implementación

La arquitectura de un sistema de DM está conformada por un conjunto de bases de datos relacionales donde residen datos numéricos, gráficos vectorizados, imágenes vectorizadas, imágenes rasterizadas, información en texto, información rasterizada, etc. La coordinación de

los generaron, debiendo preservarse para las permanentes actualizaciones a medida que surgen nuevos datos.

La producción histórica por pozo, y por consiguiente por yacimiento, está estandarizada en un alto nivel en todas las empresas del país según el modelo conocido como Capítulo IV generado en la YPF estatal. Esta información es, a su vez, requerida mensualmente por la Secretaría de Energía, hecho que ha obligado a un esfuerzo de estandarización por parte de las productoras.

La información contenida en imágenes satelitales se encuentra totalmente en forma digital, no así aquella de tipo cultural como son caminos, ríos, ciudades, etc., la cual en gran proporción se encuentra sobre papel. Muchas empresas han concentrado ambas informaciones en los llamados Sistemas de Información Geográfica (GIS); para ello han procedido a implementar un proceso masivo de escaneo de objetos gráficos.

Requerimientos del DM

Del análisis de los flujos de trabajo correspondientes a profesionales de las geociencias en diferentes empresas, previo a la implantación de un sistema de DM, se ha determinado que un 60% del tiempo total de un determinado proyecto es destinado a la búsqueda de la información y a la adecuación de la misma para su utilización en las diferentes herramientas especializadas. La disminu-

ción de este tiempo a valores mínimos constituye uno de los requerimientos esenciales para estos sistemas.

Deberá facilitarse la selección de los datos en forma gráfica a través de mapas, imágenes, secciones sísmicas o geológicas, etc., datos que una vez seleccionados podrán ser transferidos a cualquiera de las aplicaciones específicas y en el formato que éstas requieran. Las principales empresas productoras de *software* geofísico han llevado a cabo en colaboración el desarrollo de la aplicación GEOSHARE que permite el intercambio de datos entre los diferentes sistemas existentes en el mercado.

En lo referente a los prospectos sísmicos, los datos de navegación desplegados gráficamente sobre un mapa serán la vía de acceso a las secciones sísmicas y sus respectivos procesamientos e interpretaciones.

A los efectos de acceder a todo tipo de documento escrito el DM deberá contener un sistema integral de gestión documental que incluya tanto los documentos generados electrónicamente como aquellos que han sido previamente escaneados. La indexación de los documentos deberá contener varias vías de entrada a fin de facilitar la búsqueda. Para el caso de grandes volúmenes de información histórica un posible medio de búsqueda es mediante el reconocimiento de formas gráficas (*pattern recognition*) para lo cual existen productos especializados muy eficientes.

Si bien dentro de nuestro país se en-

estas bases de datos es llevada a cabo desde un módulo supervisor que a su vez ejerce la función de ventana de acceso al sistema por parte del usuario.

La ventana de acceso debe ser totalmente gráfica, con un sistema de selección y arrastre de la información requerida siendo su funcionalidad similar a un Sistema de Información Geográfica (GIS).

El ingreso de la información debe ser llevada a cabo por medio de módulos cargadores condicionados a admitir la amplia gama de formatos en que está registrada la información (perfiles, líneas sísmicas, gráficos, mapas, etc.).

El sistema de gestión documental deberá estar totalmente integrado al DM, ingresando a través de la misma ventana al que se accede para el resto de los datos. Esta condición es esencial para lograr un óptimo tiempo de acceso teniendo en cuenta el considerable volumen de información histórica registrada en papel.

Existen en el mercado desarrollos de DM integrales producidos por empresas de excelencia del ámbito petrolero. Estos productos necesitan un intensivo trabajo de adaptación a las condiciones locales de la información, que en general varían de empresa a empresa.

Una alternativa a la implementación es la provisión de uno de los tantos *software* comerciales de base de datos y proceder a desarrollar internamente su adaptación y sus módulos de acceso y de carga.

Los sistemas de gestión documental no están generalmente incluidos dentro de los sistemas comerciales de DM. En estos casos es necesario el desarrollo de una interfase de comunicación entre ambos sistemas.

El continuo avance de Internet está permitiendo sistemas de DM residentes en *software* remotos accediéndose mediante páginas WEB a través de los navegadores estándar. La velocidad de acceso remoto es el factor limitante en estos sistemas.

De contar la empresa con una Intranet se puede optar por un esquema similar al de Internet, accediéndose directamente a través de los navegadores de *windows* con un sistema cliente-servidor. Al permitir la Intranet velocidades de transmisión muy altas, este sistema es muy apropiado donde hay una gran dispersión geográfica de usuarios.

Éstos acceden a la información desde su misma PC de trabajo.

El éxito de un sistema de DM reside en parte en la calidad de los recursos humanos dedicados al mismo. Es imprescindible contar con un administrador general del sistema quien asegurará la integridad de la información y su actualización continua.

Asistiendo al administrador se deberán considerar dos técnicos con amplia experiencia en el manejo y carga de información sísmica, geológica y de perfiles. El sistema de gestión documental necesita de la dedicación de un técnico con formación especial en el tema.

A los efectos de asegurar la actualización continua de la información, sea ésta nueva u originada en proyectos, deberá establecerse un sistema de flujo de datos desde su creación hasta su ingreso al DM. Esto obliga fundamentalmente a un control estricto sobre la creación de un nuevo pozo o registración de una nueva línea sísmica.

En lo referente al *hardware* las bases de datos en general requieren de servidores UNIX con gran capacidad de almacenamiento. El acceso a la información se logra ya sea desde *workstation*, PC con emulador de Xterminal, o de contarse con una red Intranet bajo el esquema PC/emulador.

Una estimación de la dimensión del volumen de almacenamiento en disco rígido se puede obtener partiendo de casos históricos. Para yacimientos de la cuenca del golfo San Jorge y de la cuenca Cuyana se tienen los siguientes valores:

- Un legajo de pozo típico con 100 hojas y sin perfiles ocupa 10 Mb, debiéndose estimar 1Mb más por cada copia de perfil escala 1/200 escaneada.
- Una línea sísmica con *stack*, migración, y partes de observador, calculista y topografía insume 18 Mb. Con gráficos de refracción ocupa 7 Mb suplementarios.

- Una línea sísmica 2D de 10 km y 4 seg. de registración ocupa 10 Mb.
- Un cubo sísmico de 300 a 600 filas por 500 a 1000 columnas ocupa entre 4 a 8 Gb.

En la figura 1 se tiene un esquema de un sistema de *Data Management* integral para toda una empresa típica de Exploración y Producción.

Conclusiones

El escenario típico a las empresas de hidrocarburos es complejo en cuanto a la variedad de tipos de información, los medios que la contienen y el formato en que fue registrada.

Un sistema de *Data Management* debe disminuir apreciablemente los tiempos de búsqueda y captura de la información requerida en la elaboración de proyectos. El geocientista al liberarse de este tiempo muerto podrá dedicarse íntegramente al análisis propio del proyecto en cuestión.

Existen en el mercado productos integrales producidos por compañías de servicios de máxima calidad que permiten implementar un sistema de DM. Estos productos necesariamente exigen un proceso de adaptación a las condiciones particulares de cada empresa productora.

Todo sistema de DM deberá incluir una estructura de gestión documental para soportar los grandes volúmenes de información existente en medios no digitales.

El dimensionamiento del *hardware* dependerá del volumen de información existente, en función a su vez de la antigüedad de la empresa, del número de pozos, actividad exploratoria, etc.

Un sistema de *Data Management* es imprescindible para poder acceder a las nuevas y sofisticadas herramientas de interpretación y modelado. Éstas logran sus funcionalidades merced al uso masivo de todo tipo de información técnica. ●

Jorge Albano es físico egresado de la Universidad Nacional de La Plata con el título de Licenciado en Física (1964), doctorándose en la Universidad de París en 1968.

Hasta el año 1977 se desempeñó como investigador científico del CONICET, ejerciendo durante los años 1973/74 las funciones de director de la actual Facultad de Astronomía y Geofísica de la Universidad de La Plata. A partir de octubre de 1977 orienta su actividad profesional a actividades de tecnología aplicada en la industria del Petróleo, desempeñándose en ASTRA Exploración y Producción como Jefe del departamento de Ingeniería Especializada e Informática, y luego en REPSOL-YPF como responsable de Geociencias dentro de la dirección de Tecnología Informática del upstream. Fue miembro del directorio de la Comisión Nacional de Investigaciones Espaciales (CONAE), ocupándose prioritariamente de promover el uso de imágenes satelitales en la industria de los hidrocarburos. En la actualidad se desempeña como consultor independiente.