

ARQUITECTURA Y DATOS EN EL PLAN NACIONAL DE TELEDETECCIÓN

E. Domenech (*), G. Villa (*), J. J. Peces (*), J.A. Tejeiro (*) y N. Plaza (*).

(* *Instituto Geográfico Nacional, General Ibáñez de Ibero 3. Madrid. edomenech@fomento.es*

RESUMEN

Entre los objetivos del Plan Nacional de Teledetección (PNT) está el facilitar la utilización de imágenes de satélite y los datos derivados así como la disminución de los plazos existentes entre la captación y la disponibilidad de la información. Para llevar a cabo estos objetivos, es necesario diseñar un sistema de procesamiento, archivo y diseminación "on line" con servidores y redes de capacidad adecuada, que soporten diferentes niveles de procesamiento y de servicio. Son fundamentales algunos aspectos como el diseño de una infraestructura de almacenamiento adecuada al volumen de información, un flujo de trabajo adecuado, un control de la distribución y una difusión eficiente.

ABSTRACT

One of the National Plan of Remote Sensing (PNT) objectives is making the use of satellite imagery and derived data easier as well as reducing the time between the collection of data and the moment the information is available. To achieve these objectives, it is necessary to design a system for processing, archiving and online spreading using servers and networks, these servers and networks must provide suitable capacity and support different levels of processing and service. There are some fundamental aspects such as designing a storage infrastructure suitable to the volume of information, an appropriate workflow, distribution control and an efficient spreading.

Palabras clave: teledetección, almacenamiento, distribución, normalización.

INTRODUCCIÓN

El Plan Nacional de Teledetección se enmarca dentro de la primera fase del Plan Nacional de Observación del Territorio (PNOT), la fase de captura y procesamiento de imágenes aeroespaciales. El objetivo del PNT es coordinar la obtención y el tratamiento de coberturas de imágenes de satélite sobre el territorio español. De modo que se capturen y procesen una sola vez y sean accesibles por todas las Administraciones Públicas y por el resto de agentes sociales que las demanden de acuerdo con la política de datos consensuada previamente con los diferentes agentes implicados.

ALCANCE DEL PROYECTO

Las coberturas del territorio español consideradas en el proyecto se estructuran en tres niveles de resolución espacial y temporal: alta, media y baja.

Cada una de las cuales cuenta con su correspondiente flujo de datos y con la definición de los productos a generar. El resultado será un elevado número de imágenes que necesitará disponer de gran cantidad de espacio de almacenamiento. Para la cobertura de alta resolución, se requiere un espacio

de almacenamiento anual de 5,7 Tb para las más de 4.000 imágenes disponibles a este nivel (Tabla 1).

Tabla 1.- Productos de alta resolución generados en el PNT.

FASE ALTA RESOLUCIÓN	ESPACIO (Gb)	NÚMERO FICHEROS
Escenas originales	184	668
Escenas georeferenciadas	455	1002
Escenas fusionadas	1369	334
Escenas equilibradas	1369	334
Escenas color natural	1002	334
Escenas realzadas	1282	1670
Mosaico de escenas	44	4

En media resolución, se cuenta con más de 8.000 imágenes, lo que supone un requerimiento de espacio de almacenamiento de 2,1 Tb anuales (Tabla 2).

Tabla 2.- Productos de media resolución generados en el PNT.

FASE MEDIA RESOLUCIÓN	ESPACIO (Gb)	NÚMERO FICHEROS
Escenas originales	306	1020
Escenas corregidas	663	1020
Escenas en reflectividades	663	1020
Parámetro biofísico	474	5100

Finalmente, para la cobertura de baja resolución, el espacio anual que se requiere es de 1,2 Tb para las cerca de 5.000 imágenes disponibles (Tabla 3).

Tabla 3.- Productos de baja resolución generados en el PNT.

FASE BAJA RESOLUCIÓN	ESPACIO (Gb)	NÚMERO FICHEROS
Escenas originales	602	1825
Escenas en reflectividades	602	1825
Parámetros biofísicos	39	1300

Por tanto el almacenamiento de toda esta información generada en el marco del PNT supone un total de 9 Tb al año para las más de 17.000 imágenes. A esta estimación hay que añadir la correspondiente a la adquisición de las imágenes Landsat anteriores al inicio de este proyecto, esto es, las correspondientes al denominado PNT histórico, que supone adquirir todas las imágenes Landsat capturadas desde su lanzamiento con los diferentes sensores (MSS, TM y ETM+) (Figura 1), un total de 35 años de imágenes Landsat cuyo almacenamiento estimado es del orden de 40 Tb.

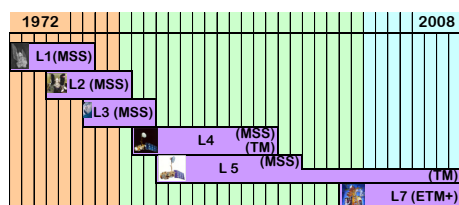


Figura 1.- Periodos de captura de sensores Landsat.

Además, hay que tener presente la posibilidad de incorporar al proyecto coberturas procedentes de otros satélites, lo que supondrá el consiguiente incremento en el espacio de almacenamiento requerido para el futuro. Teniendo en cuenta todas estas posibilidades, se presenta la estimación al alza de las necesidades de espacio de almacenamiento futuro de toda la información generada dentro del marco del PNT (Tabla 4).

Tabla 4.- Estimación de almacenamiento requerido en el PNT.

	ESPACIO NECESARIO (Tb)							
	2005	2006	2007	2008	2009	2010	2011	2012
ALTA resolución	6	3	3	6	6	6	36	36
MEDIA resolución				2	2	17	17	30
MEDIA resolución HISTÓRICA						20	20	
BAJA resolución					1	1	1	1
ESPACIO TOTAL ANUAL	6	3	3	9	9	44	74	67

REQUERIMIENTOS DEL SISTEMA

El flujo de la información de este proyecto se estructura en varios niveles cada uno de ellos con diferentes privilegios de acceso, estos niveles están formados por los centros de recepción de imágenes, los centros de proceso, los Ministerios y comunidades autónomas y el resto de usuarios finales.

El INTA, como centro de recepción, es el encargado de recibir las imágenes y enviarlas a los centros de proceso. En cada nodo de proceso se realiza el tratamiento adecuado a las imágenes para obtener los diferentes productos definidos en el PNT, mientras que se procede, en tiempo real, a la distribución de la información a los diferentes usuarios. Para esto, es necesario un sistema de archivo y diseminación “on line” con servidores y redes de capacidad adecuada que facilite la utilización de imágenes y datos derivados y que permita la gestión y el tratamiento de la información proveniente de los centros receptores y de proceso.

La estandarización de cada uno de los procesos, productos y sistemas de almacenamiento, gestión y distribución es imprescindible en el desarrollo del proyecto.

Teniendo en cuenta todo lo expuesto, el sistema a implementar para el almacenamiento y distribución de los datos del PNT, debe cumplir con los siguientes requerimientos:

- Requerimiento 1. Almacenamiento adecuado de toda la información: organizada y accesible.
- Requerimiento 2. Distribución eficiente de acuerdo con las prioridades de acceso.
- Requerimiento 3. Control de la distribución de la información acorde con la política de datos.
- Requerimiento 4. Normalización.

DEFINICIÓN DEL SISTEMA

Son diversos los aspectos a tener en cuenta para cumplir con los requerimientos marcados en el PNT en materia de almacenamiento y distribución: hay que contar con un ancho de banda adecuado al tamaño de los ficheros y al número de usuarios conectados de manera que el tiempo de respuesta sea aceptable, para lo que se deberán establecer los mecanismos adecuados con el fin de evitar tiempos excesivos de descarga de datos. A continuación se muestra la estimación del tiempo de respuesta para la descarga de datos en función del número de usuarios conectados de forma concurrente. Estos usuarios corresponderán en mayor medida a los

centros de proceso, comunidades autónomas y Ministerios (Tablas 5 y 6).

Tabla 5.- Estimación tiempo de respuesta para la descarga de datos por 1 usuario (Muñoz 2008).

	Velocidad de transmisión (Mbps)		
	50	100	1000
1 Fichero			
AR(11 Gb)	29 min	15 min	2 min
MR(0,7 Gb)	2 min	52 seg	5 seg
BR(0,3 Gb)	53 seg	26 seg	3 seg
1 Producto			
AR(1,3 Tb)	2 días	1 día	3 horas
MR(33 Gb)	2 horas	44 min	4 min
BR(1,7 Gb)	4 min	2 min	13 seg

Tabla 6.- Estimación tiempo de respuesta para la descarga de datos por 25 usuarios concurrentes (Muñoz 2008).

	Velocidad de transmisión (Mbps)		
	50	100	1000
1 Fichero			
AR(11 Gb)	12 horas	6 horas	37 min
MR(0,7 Gb)	43 min	26 min	2 min
BR(0,3 Gb)	22 min	11 min	1 min
1 Producto			
AR(1,3 Tb)	60 días	30 días	3 días
MR(33 Gb)	2 días	18 horas	2 horas
BR(1,7 Gb)	2 horas	55 min	6 min

Para evitar tiempos excesivos en la descarga de los datos, es necesario establecer los mecanismos adecuados que limiten la conexión simultánea de estos usuarios.

Para cumplir con el primer requisito, un almacenamiento adecuado, se cuenta con una EVA formada por un array de discos con capacidad suficiente para almacenar la información actual y que tiene la posibilidad de ampliar la capacidad de almacenamiento necesaria en el futuro. Mediante la implementación de un gestor documental al proyecto se proporciona una herramienta para gestionar la información, para la definición de búsquedas basadas en metadatos y para la organización de la información, permitiendo la posibilidad de incorporar un sistema de gestión de procesos (Figura 2).

El gestor documental esta basado en un repositorio único de contenidos, los organiza, gestiona su ciclo de vida y facilita un entorno de trabajo basado en la colaboración. La seguridad se implementa mediante una matriz de roles/permisos que facilitará el cumplimiento de los requisitos de distribución.

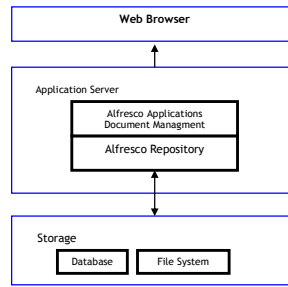


Figura 2.- Arquitectura lógica del gestor documental (Muñoz 2008).

En lo que respecta a la difusión de la información, segundo requisito, hay que establecer redes y protocolos de aplicación adecuados para la descarga de un gran número de archivos de tamaño considerable. La red IRIS (red española para la interconexión de los recursos informáticos de las universidades y centros de investigación) y la red SARA (plataforma de comunicaciones para el intercambio de datos entre unidades administrativas de cualquier Administración Pública) se presentan como dos opciones a considerar en el proyecto. Mientras que los protocolos de aplicación para la descarga de datos considerados están basados en el modelo cliente/servidor (Figura 3) y en el modelo P2P (Figura 4). En el modelo primer modelo, una aplicación en el cliente realiza peticiones a otra aplicación en el servidor (FTP, Http, WebDAV, NFS), y en el modelo P2P se basa en una red de nodos que actúan simultáneamente como clientes y servidores con respecto al resto de nodos (Direct Connet, Bit Torrent). En la selección del protocolo más adecuado, se ha considerado la eficiencia en el uso del ancho de banda, la posibilidad de limitar el número de conexiones, la fiabilidad y el soporte por parte del gestor documental.

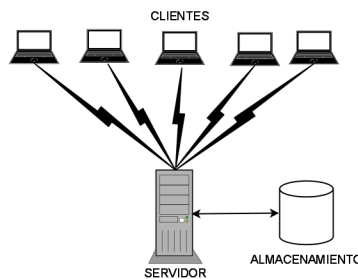
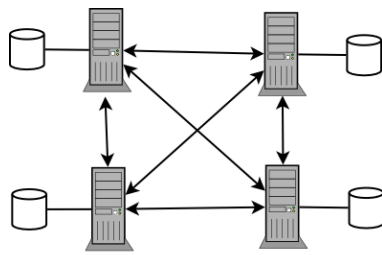


Figura 3.- Modelo cliente/servidor (Rodríguez 2008).



RED P2P PURA
Figura 4.- Modelo P2P (Rodríguez 2008).

En lo referente a la distribución de la información, hay que considerar los distintos escenarios en el PNT, que corresponden con diferentes privilegios de acceso:

- El definido por los diferentes centros de proceso de datos, en los que se debe dar tiempos de respuesta mínimos a un conjunto determinado de datos.
- El definido por los agentes participantes del PNT mediante los correspondientes convenios de colaboración, donde se proporcionarán tiempos de respuesta mínimos para toda la información del proyecto.
- El definido por el resto de usuarios que accederán a determinados productos del PNT de acuerdo con la política de datos establecida para cada uno de ellos.

En estos momentos, la solución adoptada esta basada en un servicio de descarga de archivos mediante protocolo FTP con control de usuario, password e IP.

Mediante formularios de registro de usuarios, se cumple con el tercer requisito planteado en el proyecto, el control de la distribución de los datos. Mediante estos formularios se tiene un control de qué usuarios tienen acceso a qué datos del proyecto, dado que la política de difusión de datos establece unos requerimientos de accesibilidad que deben prevalecer.

Para cumplir con el último requerimiento del PNT en materia de almacenamiento y distribución de la información, se consideran, entre otros aspectos normativos, las reglas de implementación de la Directiva Inspire, las Normas Internacionales ISO y las especificaciones OGC. Los productos del PNT, cumpliendo con la directiva Inspire, cuentan con sus metadatos ISO. Es imprescindible la documentación para alcanzar la normalización de los procesos de producción.

GT ARQUITECTURA Y DATOS

El PNT, con el objeto de contribuir al desarrollo del proyecto, cuenta con un conjunto de grupos de trabajo (GT). En este marco de trabajo, el GT de arquitectura y datos tiene la misión de definir y establecer todos los aspectos relativos a:

- Almacenamiento y Backup.
- Sistema de Gestión de datos.
- Transmisión entre centros de proceso de datos.
- Diseminación a usuarios.
- Sitio Web.
- Asignación de tareas y Cronograma del GT.
- Formatos de datos.
- Implementación de estándares.
- Metadatos (ISO 19115-2, NEM,...).

En la última reunión de este grupo de trabajo, en julio del 2008, se hizo hincapié, entre otros aspectos, al gran volumen de datos a gestionar, a la necesidad de establecer protocolos adecuados para la transferencia de archivos y a la necesidad de contar con un gestor documental para controlar los flujos de datos de todos los productos incluidos en el proyecto. Esto es, todos los aspectos necesarios para el cumplimiento de los requisitos del PNT en materia de almacenamiento y distribución ya enunciados anteriormente.

Como consecuencia de los trabajos de este grupo de trabajo se ha implementado una página Web para impulsar la difusión de la información del proyecto.

BIBLIOGRAFÍA

Antonio Arozarena Villar, Lorenzo García Asensio, Guillermo Villa Alcázar, Emilio Domenech Tofiño. Conama 2008. *Plan Nacional de Observación del Territorio en España*.

Equipo Técnico Nacional. 2005. *Especificaciones Técnicas para el Plan Nacional de Teledetección (PNT)*. Madrid.

GT de arquitectura y datos del PNT. 2008. *Acta de la reunión*.

Instituto Geográfico Nacional. 2009. *Documento PNT versión 2.0*. Madrid.

Muñoz Cruzado, Almudena. 2008. *Estudio de soluciones de distribución masiva en el contexto del Plan Nacional de Teledetección*.