

Integración de la Infraestructura de Datos Espaciales de la Generalitat Valenciana en el nodo IDE de la Comunitat (IDECV)

PROCESO DE UNIFICACIÓN DE ARQUITECTURAS

CABEZUDO, Laura; CONTI, Luis; FDEZ. DE YBARRA, Juan; SÁNCHEZ, Fernando

La implantación y desarrollo de la Infraestructura de Datos Espaciales de la Comunitat Valenciana (IDECV) se ha venido desarrollando desde el año 2007 bajo el liderazgo del Institut Cartogràfic Valencià como órgano coordinador de la ejecución de INSPIRE en la Comunitat. La infraestructura se ha materializado en estrecha colaboración con la Dirección General de Tecnologías de la Información y Comunicación de la Generalitat.

En la actualidad, el nodo IDECV aloja la Infraestructura de Datos Espaciales corporativa de la Generalitat (IDEGVA), dada la centralización de recursos informáticos que desde hace unos años se impuso en el Gobierno Valenciano.

La intensa actividad que en la última década ha rodeado la gestión de los datos geoespaciales en la administración ha provocado el aumento de inversiones de forma no coordinada, ya que la velocidad de aparición de nuevas demandas ha sido mayor que la creación de marcos de colaboración e implantación de bases de trabajo. Por esa razón, se han creado distintas Infraestructuras de Datos Espaciales dispersas que no siempre han conseguido cumplir los objetivos para los que fueron creadas y que suponen un alto coste de mantenimiento.

Dada la situación actual, y los criterios que siempre deben prevalecer de eficacia y eficiencia dentro de la administración, la evolución de las distintas soluciones dentro del gobierno valenciano tiende a la centralización de repositorios y la adopción de soluciones tecnológicas comunes.

Por tanto, la creación de infraestructuras centralizadas, como es el caso de la IDEGVA, no sólo requiere un gran esfuerzo técnico sino un esfuerzo aún mucho mayor de divulgación y concienciación de la necesidad de transparencia de las actuaciones relacionadas con los datos geográficos, de adopción de acuerdos de acceso, uso y mantenimiento de los datos y servicios, de formación para el uso de herramientas que aseguren la armonización y del necesario empoderamiento de los órganos de decisión y coordinación para llevar a buen fin la infraestructura.

PALABRAS CLAVE

IDECV, IDEGVA, arquitectura de sistemas, centralización de bases de datos.

INTRODUCCIÓN

La Infraestructura de Datos Espaciales de la Comunitat Valenciana (IDECV) que ejecuta la implantación de la Directiva 2007/2/EC tiene su primer embrión en el año 2007 dentro del proyecto terrasit, orientado a cubrir las necesidades de la Generalitat Valenciana en materia de estrategia territorial en cuanto a recursos cartográficos se refiere. El proyecto está coordinado por el Institut Cartogràfic Valencià y en ese momento, ejecutado en su totalidad por personal propio del organismo.

El proyecto terrasit no tiene como objetivo únicamente la creación de una IDE dentro de los términos de INSPIRE, trabaja dentro de términos más amplios dando a conocer cualquier información geoespacial de uso público o privado de la administración basándose en servicios interoperables.

El proyecto terrasit en el año 2011 ya había creado una red dispersa de proveedores de datos y servicios dentro de la Generalitat. En esas fechas, la política de la Generalitat en cuanto a

tecnologías de información y comunicación cambia hacia un modelo centralizado. Ese cambio ha ocasionado importantes problemas al desarrollo y coordinación de la infraestructura, pero finalmente lo que parecía ser un obstáculo puede convertirse en el punto de inflexión que potencie el proyecto.

El objetivo de este artículo, es presentar el actual proceso de transformación de la infraestructura para adaptarse a la nueva arquitectura y la redefinición de roles, en un cambio tan importante como es la reestructuración de todas las infraestructuras de sistemas de información dispersos en los diversos organismos y entidades de la Generalitat.

ARQUITECTURA INICIAL DE LA INFRAESTRUCTURA DE DATOS ESPACIALES DE LA COMUNITAT VALENCIANA Y SU RELACIÓN CON LA INFRAESTRUCTURA DE LA GENERALITAT

La arquitectura de la IDECV se basa en un nodo central de divulgación y acceso a los datos que son proveídos por los diferentes generadores de información geoespacial.

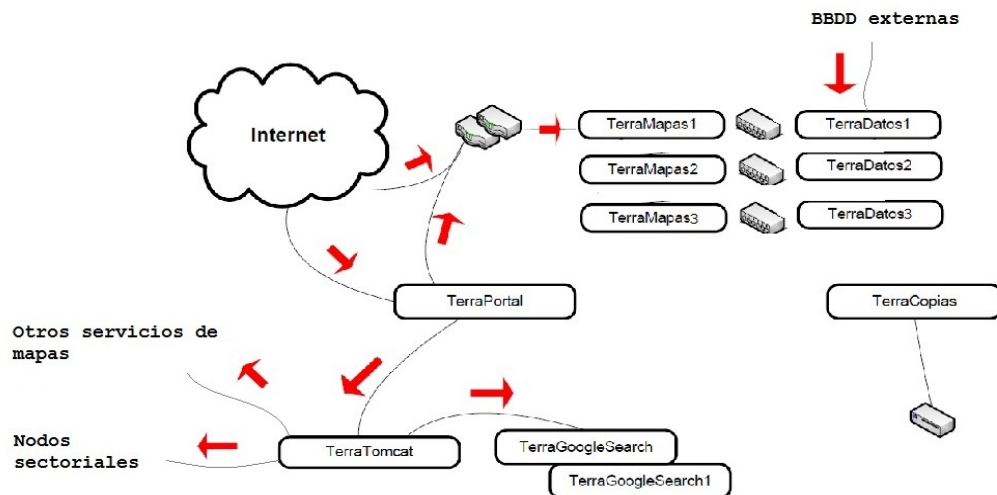


Figura 1: Arquitectura de la IDECV en el año 2011

El nodo central en su plataforma de producción se localiza en el CETESI (Centro de Telecomunicaciones y Sistemas de Información) con un servicio de mantenimiento 24x7, mientras que la plataforma de desarrollo y preproducción se encuentra en el rack de servidores del ICV.

Este nodo se compone de geoportal (implementado sobre Drupal), catálogo de metadatos (Geonetwork), buscador *Google Appliance*, tres servidores de mapas (Mapserver) con BD PostGis replicadas y dos balanceadores de carga que distribuyen las peticiones entre los servidores. Conectados a este nodo central a través del catálogo existe una red distribuida de servidores de publicación de servicios.

Como se muestra en la figura 1 el acceso a la información geoespacial se realiza a través del frontal terrasit-IDECV desde donde se accede a todos los servicios dados de alta en el nodo accediendo directamente a los servidores de publicación que están alojados, o bien en el nodo o bien en otros servidores de publicación externos.

Los servicios creados por el equipo de proyecto del ICV pueden ser distribuidos o centralizados, en función de la capacidad técnica en materia de publicación y manipulación de datos geoespaciales y

densidad de capas producidas que tenga cada organismo productor. Podemos encontrar tres escenarios:

a) Se trata de un organismo de alto nivel de producción de datos geoespaciales que además dispone de personal técnico capacitado para crear un nodo sectorial. Este es el caso de la Conselleria competente en materia medioambiental.

b) Se trata de un organismo que potencialmente dispone de mucha información, aunque en la actualidad no dispone de personal que específicamente esté preparado para la creación de un nodo sectorial, o bien se trata de organismos que trabajan con información sensible y no desean que sus BB. DD. sean atacadas desde un nodo externo. En estos casos se crea en los servidores del organismo una plataforma de publicación compuesta por una BB. DD. y un servidor de mapas dentro de la misma máquina, donde la BB. DD. de publicación está alimentada por la BD funcional por mediación de un *script* de actualización de datos.

c) Se trata de organismos que dentro de sus funciones crean BB. DD. con referencias geoespaciales, casi siempre implícitas. En este caso, se utiliza la plataforma de publicación del nodo y se actualiza invocando remotamente la BD origen.

Los roles dentro de esta arquitectura son:

Rol del organismo responsable de la infraestructura:

- Definición de la arquitectura de sistemas IDE (*software, hardware, comunicaciones*)

- Tareas para la búsqueda de datos espaciales

- Acuerdos y coordinación de tareas para el acceso y publicación de los datos

- Implantación y mantenimiento de infraestructuras externas

- Georreferenciación de BB. DD., creación de *scripts* y servicios de mantenimiento de la BD de publicación y creación de servicios de mapas

- Formación para el uso de la infraestructura

Rol del organismo proveedor de datos:

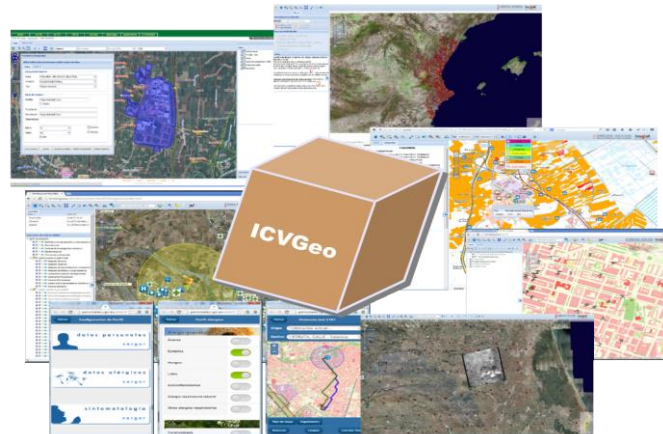
- Acuerdos sobre la definición y características de la información a publicar

- Mantenimiento de la información con garantías de calidad (compleción, fiabilidad temporal, exactitud posicional y de atributos...)

- Asegurar el acceso a los datos desde el nodo

- Asegurar la formación de personal para el uso de servicios asociados a la infraestructura

La mayor parte de los acuerdos para la publicación de datos llevan implícito el desarrollo de



aplicaciones para la explotación de los servicios.

Figura 2 : Aplicaciones generadas con ICVGeo

Con respecto a los servicios publicados, en el año 2011 son en su mayoría servicios WMS, encontrándonos algunos WFS-T creados para edición en aplicaciones a terceros, algunos WFS no publicitados y WMTS creados principalmente en series cartográficas de gran extensión que son utilizadas como servicios básicos para visualizadores.

Tomando como referencia los servicios WMS publicados a través del servicio de mapas del nodo autonómico, podemos observar que en su mayoría son datos generados por la Generalitat, de estos el 93 % presentan datos cartográficos del ICV y únicamente un 6 % son referentes a otras competencias de la GVA (no se incluyen los servicios en materia medioambiental puesto que se publican desde su propio nodo sectorial). Véase la figura 3.

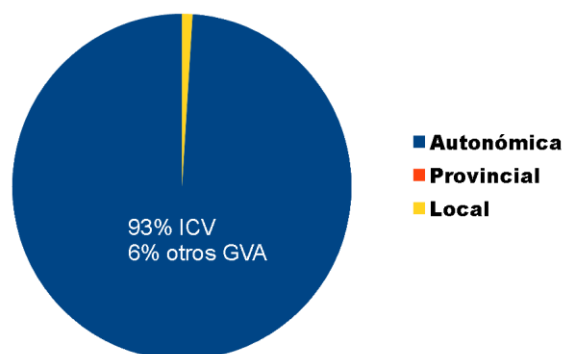


Figura 3 : Distribución de servicios WMS publicados por el nodo en función de la administración propietaria de los datos espaciales

Dado que tanto el personal como los recursos informáticos pertenecen a la GVA y están presentes en su presupuesto anual, no tiene sentido hablar de la creación de una nueva Infraestructura de Datos Espaciales de la GVA puesto que *de facto* ya existe..

Pros y contras del escenario original (IDECV 2011):

- No existe regulación legislativa de la Comunitat Valenciana que promueva la implantación de las IDE en la Comunidad Valenciana y por tanto no se incluye como objetivo del gobierno autonómico lo que deriva en:

No existe una mesa a alto nivel que impulse las tareas de implantación, lo que dificulta la puesta en común de datos geoespaciales, especialmente aquellos bajo el ámbito de aplicación de INSPIRE.

La creación de la IDECV no se desarrolla como un proyecto horizontal sino como un proyecto del ICV cuyos recursos económicos y de personal especializado son muy limitados.

- La independencia técnica del ICV en cuanto a definición de estándares y tecnología a utilizar facilita el desarrollo interno, ya que posibilita la utilización de *software* y desarrollos sobre plataformas libres, centrandó únicamente la inversión en *hardware*.
- El desarrollo de aplicaciones para proveedores de datos dentro de la arquitectura de la IDECV resulta una sobrecarga sobre los pocos perfiles de desarrollo de los que dispone el ICV. Estas aplicaciones se desarrollan sobre un núcleo JavaScript (ICVGeo) y se crean *plugins* y *base layers* específicos para cada visor.

Normalmente estas aplicaciones suelen ser críticas, con lo que se desvían los pocos recursos existentes a estas tareas. Los plazos y recursos limitados provocan la gran proliferación de servicios WMS OGC y servicios de procesamiento que no cumplen la normativa de INSPIRE.

Posteriormente, en el año 2015 para no mermar los recursos asignados a la implantación de INSPIRE, se opta por derivar las aplicaciones *ad hoc* para proveedores de datos a una tecnología ESRI, ArcGis Online, que permite apersonas con un perfil menos especializado la creación de estos visualizadores.

- La solución genera un alto control y conocimiento interno de la infraestructura.

ESTRATEGIA TIC DE LA GENERALITAT VALENCIANA Y SU IMPACTO INICIAL EN EL NODO AUTONÓMICO

En 2011 la Dirección General de Tecnologías de la Información y las Comunicaciones de la Generalitat Valenciana asume las funciones en materia de modernización de la Administración, seguridad de la información, planificación, coordinación, autorización y control de las tecnologías de la información, las telecomunicaciones y comunicaciones corporativas, y la teleadministración de la Generalitat.

Desde la DGTIC se plantea la estrategia TIC 2011-2015, donde prevalecen las siguientes directrices:

Racionalización TIC

La Generalitat racionaliza el uso de sus recursos TIC mediante la centralización de competencias. Así, se aprovechan las ventajas del modelo centralizado para encontrar sinergias entre proyectos, identificar y eliminar recursos redundantes y obtener los ahorros propios de las economías de escala. Esto incluye, además, la implantación de medidas para la reducción generalizada de gastos corrientes en TIC.

Innovación GVA

La Generalitat Valenciana requiere la incorporación de las TIC a todos los procesos de la

administración para la innovación tecnológica de su gestión, ya que son una herramienta clave para la provisión de servicios públicos de calidad. Las TIC permiten mejorar la eficacia y la eficiencia de la administración y ofrecen información fiable y actualizada a los órganos directivos para la aplicación de sus políticas. Asimismo, esta incorporación resulta indispensable para el impulso de la administración electrónica, ya que las TIC facilitan la simplificación de procedimientos y la reducción de cargas administrativas para ciudadanos y empresas.

Sociedad digital

Las TIC constituyen un sector que destaca como motor de crecimiento sostenible en la Comunitat Valenciana, ya que aumenta la productividad de otros sectores, fomenta el empleo y favorece la cohesión social. Con la innovación tecnológica, la Generalitat dinamiza el sector TIC y, a través de su poder de compra, potencia la creación de empleo TIC y fomenta la inversión privada en infraestructuras y servicios en nuestro territorio.

Dentro de la estrategia de la DGTIC la racionalización de las TIC afecta seriamente al proyecto de implantación de la IDECV puesto que la centralización de recursos conlleva la definición de estándares y directrices comunes para los proyectos TIC de la GVA, la centralización del personal informático, y la tramitación de cualquier expediente cuyo objeto entre dentro de las competencias de la DGTIC como adquisición y mantenimientos de *software* y *hardware*. Como consecuencia, dado que hasta ese momento todo el proyecto se ha llevado a cabo dentro del ICV, la racionalización afecta al proyecto en los siguientes puntos:

1. Los estándares definidos por el servicio de calidad de la DGTIC no están alineados con los ya implantados en el proyecto, por lo que incumple la infraestructura los requisitos para que sea aceptada por la DGTIC y no puede ser reutilizada la infraestructura ni el conocimiento adquirido por el personal del proyecto.
2. Al no cumplir la infraestructura los estándares requeridos se proponen nuevos proyectos que duplican la infraestructura ya existente.
3. La Infraestructura de Datos Espaciales se plantea como un proyecto TIC enteramente informático por lo que debe pasar en su totalidad, tanto en los sistemas como en la parte de definición de servicios de mapas y otras componentes geográficas, a equipos de la DGTIC, lo que conllevaría una desvinculación del organismo coordinador e impulsor de la implantación de INSPIRE en la Comunitat Valenciana con la implantación y mantenimiento del proyecto que la ejecuta.
4. Se centralizan los recursos de personal informático, incluyendo aquellos integrantes del equipo de desarrollo e implantación IDE correspondientes a sistemas informáticos.

El resultado inicial es una dicotomía de competencias geográficas e informáticas a nivel técnico y ejecutivo del proyecto que, como veremos a continuación, se solventa con la creación de un equipo mixto.

ADAPTACIÓN DE LA ARQUITECTURA IDECV ORIGINAL A LA NUEVA ESTRUCTURA INFORMÁTICA

La nueva estructura de la DGTIC requiere una modificación interna de los sistemas de la IDECV ya que se ha de adaptar a una arquitectura de tres niveles, especialización de la arquitectura cliente-servidor, en la que la carga se divide en tres partes (o capas) con un reparto claro de funciones: una capa para la presentación (interfaz de usuario), otra para el cálculo (donde se encuentra modelado el negocio) y otra para el almacenamiento (persistencia). Una capa solamente tiene relación con la siguiente. Esta arquitectura además de dividir las funciones que después pasarán a los equipos competentes en cada capa, le da el máximo nivel de protección a los datos. que no son accesibles directamente desde Internet sino a través de la capa de aplicaciones. En la figura 4 se representan las tres capas de la IDECV.

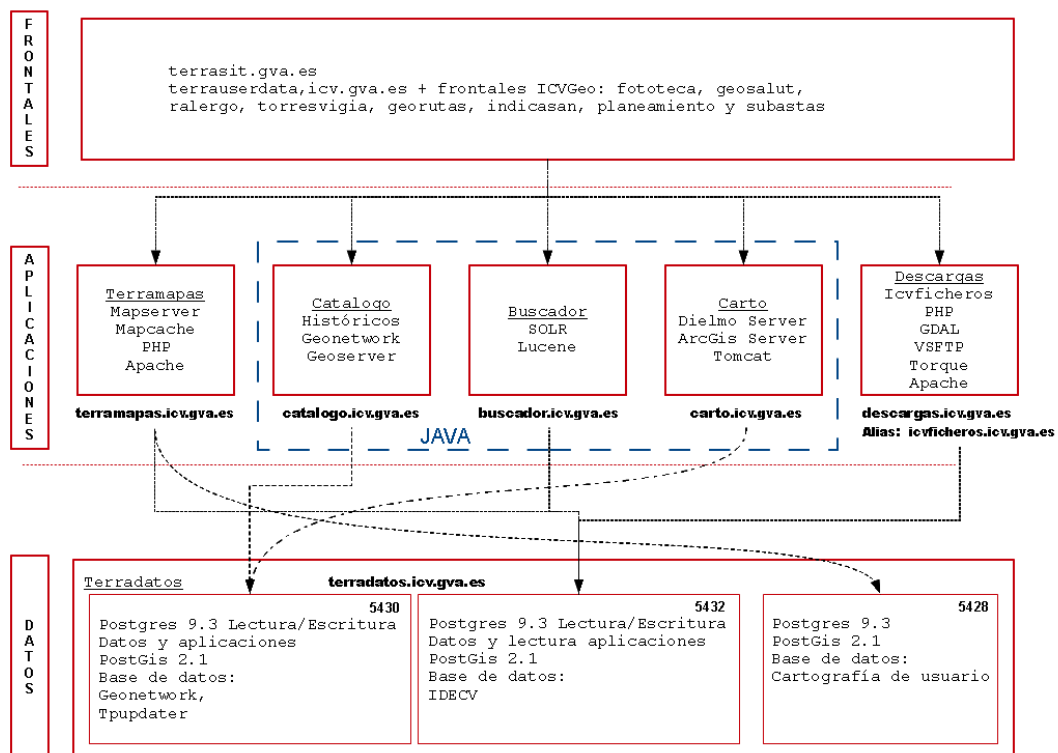


Figura 4: Arquitectura de tres niveles IDECV

En la capa 1 encontramos el geoportal de la IDECV y la clase ICVGeo como una extensión de GeoExt, Ext JS y OpenLayers, así como todos los *plugins* de visualizadores para terceros, creados y mantenidos por la infraestructura.

En la capa 2 se encuentran los servidores de publicación de servicios y metadatos, servidor de descargas, servidor de geoprocetos y buscador de contenidos. Podemos ver que se ha migrado el buscador de tecnología Google a tecnología Apache Lucene y se dispone en la actualidad de un buscador SOLR. A diferencia de la configuración anterior se han separado en dos máquinas el catálogo y el buscador por razones de rendimiento.

En la capa 3 se incluyen los datos que no son accesibles directamente desde Internet. A través del puerto 5428 y 5430 accedemos a la BD de Geonetwork, cartografía subida por los usuarios de terrasit, BD resultado de aplicaciones a terceros, BD de descargas y otras necesarias para el funcionamiento de la plataforma. A través del puerto 5432 accedemos a la BD que contiene los datos espaciales que se publican desde el servidor de mapas del nodo. Todos los datos están contenidos en un mismo Postgres en terradataos.icv.gva.es.

Fuera de esta arquitectura de tres capas existe un *pool* de almacenamiento de la GVA al que se accede desde la capa de aplicaciones, desde la ftp para la descarga de ficheros de cartografía y desde servidor de mapas para acceso a los ficheros .map que crean y almacenan los técnicos del servicio de proyectos del ICV.

CREACIÓN DE UN REPOSITORIO ÚNICO DE DATOS ESPACIALES DE LA GENERALITAT

En la arquitectura actual de la IDECV se han centralizado todas las BD Postgres que se encontraban distribuidas en las máquinas de terramapas, catálogo, descargas y terraportal, creando una única BD espacial del nodo dentro de NICA90 mantenida por un equipo multidisciplinar DG TIC-ICV.

Paralelamente, la centralización de recursos conlleva la centralización de frontales, aplicaciones y datos de la GVA. Se formaliza de esta manera la integración de la IDEGVA en la IDECV a través de un

proyecto que se inicia con la creación de un repositorio único de datos espaciales donde se incluyen las BD espaciales de publicación, hasta ahora dispersas en la administración autonómica. A partir de los datos del repositorio centralizado se crearán los servicios de mapas mediante los servidores de publicación de la infraestructura, eliminando de esta forma los recursos redundantes gestionados por una misma administración.

El proyecto a su vez se ve reforzado al integrarse el grupo de trabajo que mantiene la IDE sectorial de datos medioambientales funcionalmente en el ICV, con lo que pasao a formar parte del equipo mixto DGTIC-ICV.

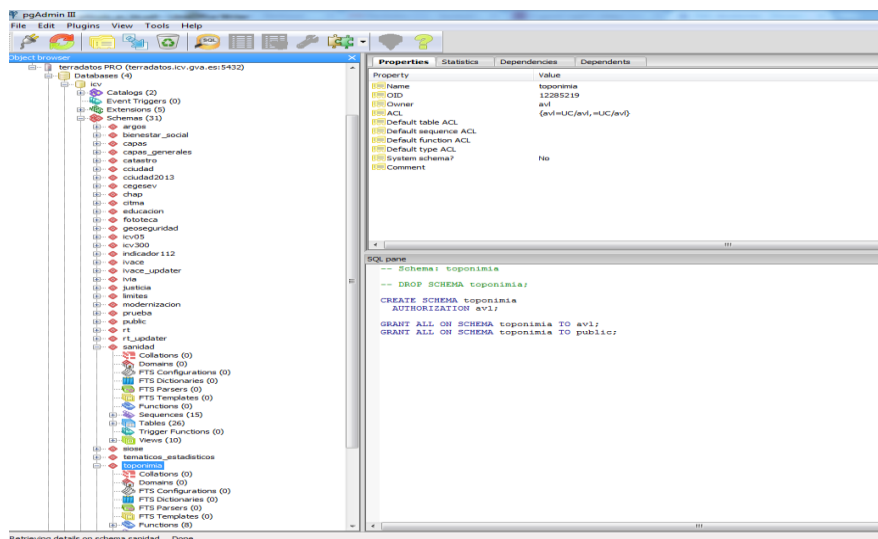
El proyecto de integración de la IDEGVA tiene la siguiente hoja de ruta:

Fase 1.- Crear una única BD Postgres alimentada con los datos de publicación de la GVA. En este punto, una vez unificadas los Postgres del nodo IDECV, se incluirán todas las BD espaciales de publicación dispersas en la administración autonómica iniciando los trabajos con aquellas creadas por el equipo de proyecto IDECV antes del 2011. Estos grupos de datos se publican con Mapserver por lo que directamente pasarán a publicarse con el servidor de mapas de NICA90. Este es el caso de la Conselleria de Educación, o el Centro de Gestión y Seguridad vial de la Generalitat.

Posteriormente se iniciarán los trabajos a nivel de BD con nodos IDE sectoriales de la Generalitat. Tal es el caso de la Conselleria de Sanidad y la Conselleria de Agricultura, Medio Ambiente, Cambio Climático y Desarrollo Rural.

En la actualidad la BD Postgres centralizada de IDECV se conforma según esquemas que representan el propietario de los datos o producto.

Figura 5:
Postgres



BD

centralizada de publicación

Para la actualización de las BD de publicación dejamos de utilizar PHP/Postgres sustituyéndolo por funcionalidades de conectores DBLink de Postgres.

Fase 2.- Integrar los servicios de publicación dentro de la estructura de la IDECV-IDEGVA.

Existen varias tecnologías de servidores de publicación funcionando simultáneamente en la Generalitat:

Servidor	Infraestructura	Uso
Mapserver	IDECV	WMS, WMTS, WFS

ArcGis Server	IDECV y IDECITMA	WPS, WMS, REST
Geoserver	IDESAN, IDECV	WMS, SLD, impresión

En la integración actualizamos las versiones mantenidas por la DGTIC, requiriendo previamente a la puesta en producción un estudio detenido de las implicaciones que esto supone en las aplicaciones clientes de los servicios.

En la integración de nodos sectoriales como en el caso de la Infraestructura de Datos Espaciales de Sanidad (IDESAN) y la Infraestructura de Datos Espaciales Medioambientales (IDECITMA) nos encontramos como se aprecia en las figuras 6, 7 y 8 con estructuras similares a la del nodo IDECV.

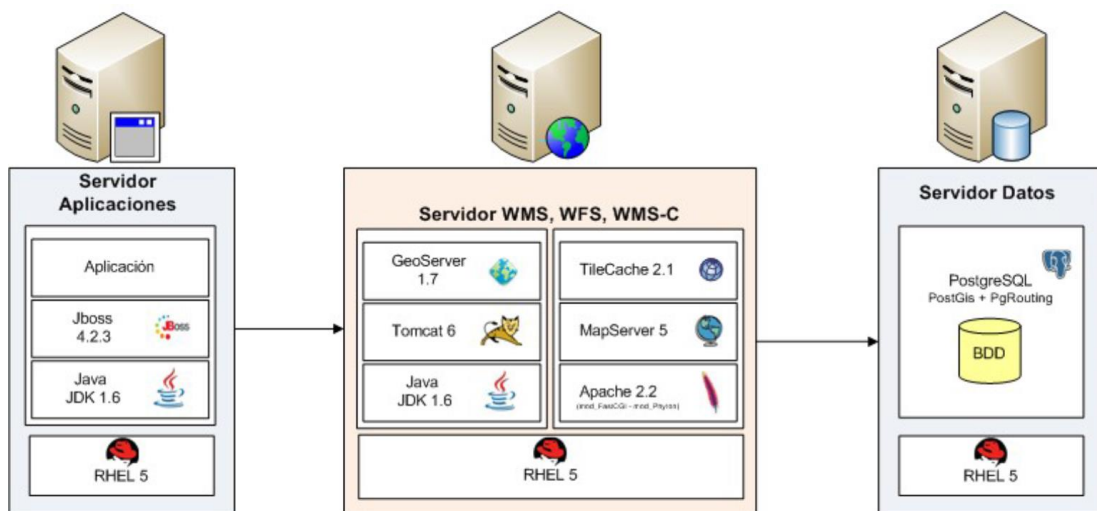


Figura 6: Arquitectura física IDESAN

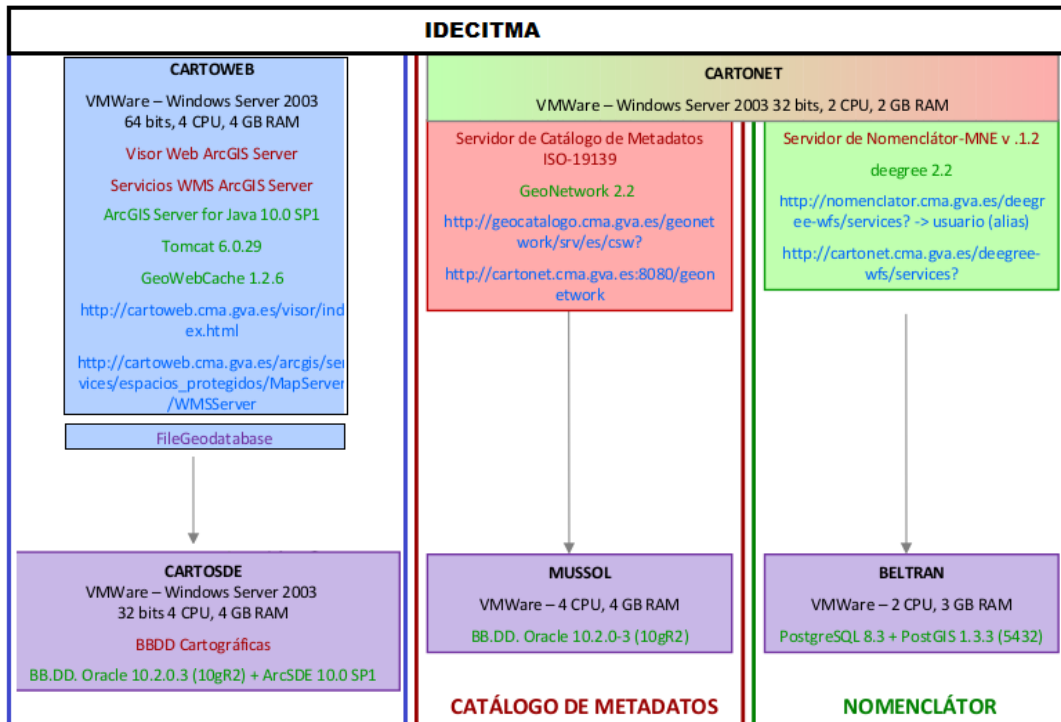


Figura 7: Arquitectura física IDECITMA

Las tres arquitecturas tienen configuraciones comunes propias de la arquitectura general de una Infraestructura de Datos espaciales. La unificación de las tres arquitecturas (IDESAN, IDECITMA e IDECV) dentro del nodo del NICA90 se aplica al nivel 2 (servidores de publicación de servicios y catálogo de metadatos) y nivel 3 de BD espaciales. Quedan por tanto fuera de la integración las aplicaciones funcionales internas o visualizadores específicos desarrollados como frontales de acceso.

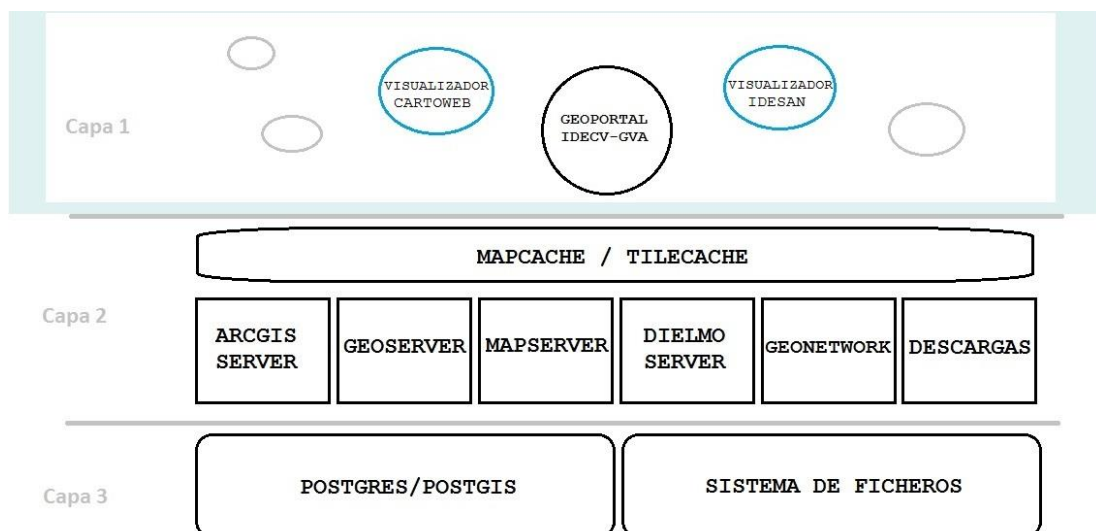


Figura 8: Esquema de arquitectura inicial de integración de infraestructuras de la GVA en el nodo de IDECV del CPD corporativo de la Generalitat

Fase 3,- Optimización de la Infraestructura.

Una vez finalizada la primera integración se plantean las siguientes posibilidades a nivel de arquitectura:

- Eliminar el servidor de descarga, para realizar ésta desde el servidor de Opendata de la Generalitat (GVA Oberta) o bien directamente desde el servidor de catálogo.
- Migrar de MapCache a MapProxy.
- Estudiar la optimización de tecnologías de publicación de servicios.
- Eliminar los *tileindex* de los ficheros *shp* o imágenes, para sustituirlos por índices espaciales de PostGis.

Para el correcto funcionamiento del nodo IDE es fundamental definir los roles de los integrantes del equipo de proyecto y aquellas actuaciones a distintos niveles que deben realizarse de forma urgente para un correcto funcionamiento del sistema y cumplimiento de la Directiva europea INSPIRE.

Para la implantación de la plataforma se han definido unas premisas:

- La Infraestructura de Datos Espaciales no contempla desarrollo de aplicaciones funcionales, es decir, todas aquellas aplicaciones clientes de los servicios IDE desarrolladas en diferentes tecnologías no son creadas ni mantenidas por el equipo de proyecto, aunque sí existe la figura del equipo como asesor de los desarrollos de componente geográfica de la GVA.
- Los datos geospaciales contenidos en la infraestructura tienen como objetivo su publicación a través de servicios y son sólo accesibles a través de ellos.
- La distribución de roles en la infraestructura está compartida entre el Institut Cartogràfic Valencià y el servicio de Sistemas de la DGTIC creando un grupo de trabajo que coordina y mantiene la geomática de la Generalitat.

En este grupo de trabajo el Institut Cartogràfic Valencià aporta perfiles de coordinación con los distintos proveedores de datos, y especialistas en geomática responsables de la creación de las infraestructuras de la Generalitat. Por parte de la DGTIC se aportan perfiles especializados en BD espaciales, servidores de aplicaciones, sistemas operativos y servicios de componente espacial. Para asegurar el mantenimiento de la infraestructura se deben realizar otras acciones a nivel técnico y político. Es fundamental el desarrollo de herramientas y definición de procesos que permitan de forma ágil la creación y mantenimiento de cada uno de los elementos que conforman la infraestructura, pero todos los trabajos y acuerdos a nivel técnico sólo conseguirán el objetivo de optimización de recursos, transparencia dentro y fuera de la administración, impulso de la sociedad de la información, mejora de la gestión de la administración y por supuesto el cumplimiento de INSPIRE, si existe una legislación autonómica que asegure su desarrollo, así como una mesa de alto nivel en la administración que asegure la dotación de medios, el cumplimiento de las normas, y el establecimiento de una política común de datos espaciales.

REFERENCIAS

- [1] PostgreSQL 9.3.14 Documentation, <https://www.postgresql.org/docs/9.3/static/dblink.html>
- [2] Estrategia TIC 2011-2015 la hoja de ruta de la Generalitat Valenciana en TIC, http://www.dgtic.gva.es/documents/85347/355059/estrategia_tic_2011-2015.pdf/7f9edffa-76b8-4dcb-b999-db57ae3a6aa4
- [3] Agencia Valenciana de Salud, Área de Informática, Telecomunicaciones y Organización: IDECS-Arquitectura Física (2010)
- [4] Tony Marston 2012 : What is the three-tier architecture?, <http://www.tonymarston.net/php-mysql/3-tier-architecture.html>

AUTORES

Laura CABEZUDO DE LA MUELA
cabezudo_lau@gva.es
Institut Cartogràfic Valencià
Servici Projectes Cartogràfics

Luis CONTI BUENO
conti_lui@gva.es
Institut Cartogràfic Valencià
Servici Projectes Cartogràfics

Juan FDEZ. DE YBARRA DEL REY
fernandezdeybarra_jua@gva.es
Consell. de Hacienda y Modelo Económico. DGTIC. Servicio de Sistemas y producción

Fernando SÁNCHEZ MIRAMÓN
sanchez_fermir@gva.es
Institut Cartogràfic Valencià
Servici Projectes Cartogràfics