*E1.2*

***Propuesta de Diseño de Base de Datos y Migración***

*v1.0*

04 Septiembre 2019

Para cualquier consulta o cuestión pendiente sobre la oferta la persona de contacto es:

|  |  |  |
| --- | --- | --- |
| **Nombre** | **Teléfono** | **e-mail** |
| José Vicente Higón Valero | +34 600 55 44 74 | [jvhigon@gvsig.com](mailto:jvhigon@gvsig.com) |

Información general:

Asociación gvSIG

CIF G98125362

Calle Ángel Guimerá 61, puerta 3

[http://www.gvsig.com](http://www.gvsig.com/)

**Control de cambios**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Versión** | **Responsable** | **Organismo** | **Descripción del cambio** | **Fecha** |
| v.1.0 | Francisco José Peñarrubia | Asoc gvSIG | Documento inicial | 07/18/19 |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**Índice**

[1. Introducción](#_gjdgxs) **3**

[2. Tablas a utilizar](#_30j0zll) **4**

[**3. Campos modificados**](#_1fob9te) **5**

[**4. Vistas e índices.**](#_3znysh7) **6**

[**5. Consideraciones Finales**](#_2et92p0) **10**

[**6. Scripts de Migración**](#_tyjcwt) **10**

# 1. Introducción

El siguiente documento hace referencia al plan de trabajo definido en el documento *Sección VI. Lista de Bienes y Servicios* de la licitación **Abreviada Nº 112/2018**:

[...]

*4. Entregables*

*Durante la ejecución del proyecto, la empresa adjudicataria deberá cumplir con*

*los siguientes entregables:*

*1) Documento de diagnóstico de Datos y Sistema Actual*

***2) Propuesta de diseño de base de datos y migración.***

*3) Plan de Trabajo*

*4) Documento de Requerimientos.*

[…]

Se trata de realizar una propuesta de cambios al esquema actual (TSUBASA) en base al análisis de la solución actual (Documento 1: Diagnóstico de Datos y Sistema Actual).

# 2. Tablas a utilizar

En base al análisis, se han detectado tablas que no son útiles para el proyecto de Callejero (se usan para otros proyectos). Esas tablas se pueden mantener o eliminar, aquí vamos a describir las que sí vamos a usar.

|  |  |
| --- | --- |
| **Tabla** | **Num. Registros Estimado** |
| alias | 37.410 |
| alias\_departamento | 1 |
| alias\_entidad\_colectiva | 1 |
| alias\_localidad\_geo | 758 |
| alias\_numero\_puerta | 2.990 |
| alias\_tramo\_ruta | 183 |
| bloque | 0 |
| calle | 26.496 |
| calle\_departamento | 26.579 |
| calle\_localidad | 47.839 |
| codigos\_clasificacion | 1.153 |
| codigos\_postales\_vigentes | 121 |
| departamento | 20 |
| entidad\_colectiva | 5 |
| localidades\_no\_oficiales | 1.916 |
| manzana | 54.647 |
| origen | 4 |
| origen\_alias | 14 |
| palabra | 9.564 |
| punto | 1.596.041 |
| punto\_notable | 1.114 |
| ruta | 115 |
| sinonimo | 896 |
| sinonimo\_entidad | 2 |
| sinonimo\_localidad | 10 |
| solar | 841.003 |
| spatial\_ref\_sys | 5.757 |
| tipo\_alias | 5 |
| tipo\_entidad\_colectiva | 13 |
| tipo\_punto\_notable | 4 |
| tipo\_tramo\_ruta | 0 |
| tipo\_vialidad | 18 |
| torre | 0 |
| tramo\_ruta | 186 |
| tramos | 6.251 |

La tabla con más registros es **punto**, y contiene una geometría de tipo punto en el sistema de referencia EPSG:32721 (WGS 84 / UTM zone 21).

Esta tabla es la que hemos utilizado como base para el piloto de búsqueda de direcciones, porque parece la más completa. Sin embargo, contiene muchos campos con valores nulos, probablemente porque contiene datos relativos a conceptos distintos. Hay puntos que provienen de las calles, otros de los puntos notables de interés, y otros de manzana y solar.

Quizás deberíamos plantearnos separarla en varias tablas, y buscar de manera separada y ofrecer los datos relevantes de cada consulta.

En el documento de análisis se especifican las consultas que hemos empleado para detectar errores e incoherencias en las tablas. Estas consultas se deberían emplear cuando se haga la parte de administración para detectar y ayudar a corregir los errores detectados.

De manera opcional, planteamos la posibilidad de dividir la tabla de punto en la tabla “portal\_pk” con los números de portal y puntos kilométricos, y “poi” con los puntos notables, inmuebles singulares, etc. Los puntos relativos a solares y manzanas se pueden separar también, o directamente eliminarlos y hacer las búsquedas sobre la tabla de solares y manzanas.

# 3. Campos modificados

Se propone eliminar los siguientes campos:

**Tabla punto:**

ordinal, ordinal\_calculado, entidad\_colectiva\_id, zona\_paquete

Si no se usan para sincronizar con las fuentes externas: forma\_de\_origen, forma\_de\_modificacion

Si se separan las tablas: solar\_id, nombre\_inmueble, tipo\_punto\_notable\_id, bloque\_id, torre\_id

**Tabla calle:**

geom, cantidad\_puntos, ruta\_id

Si no se usan para sincronizar con las fuentes externas: cod\_nombre, origen

En las tablas que se vayan a realizar mantenimiento, se propone añadir los campos siguientes: **fecha\_alta** : Fecha en que aparece el objeto geográfico en la realidad o en que se recopila de la fuente original.

Lista de Valores "20080521" (Ejemplo)

Tipo de Valor Integer

Longitud 8

**fecha\_baja** : Fecha en que un objeto geográfico deja de existir en la realidad o en que se elimina de la fuente original.

Lista de Valores "20070315" (Ejemplo)

SinValor

Tipo de Valor Integer

Longitud 8

**alta\_db** : Fecha en que se inserta un nuevo objeto geográfico en la base de datos.

Lista de Valores "20080521" (Ejemplo)

Tipo de Valor Integer

Longitud 8

**baja\_db** : Fecha en que se elimina un objeto geográfico de la base de datos

Lista de Valores "20070315" (Ejemplo)

SinValor

Tipo de Valor Integer

Longitud 8

**Campos añadidos:**

En la tabla “tramos” se propone añadir los campos:

ncarriles : Número de carriles

Tipo de Valor Integer

Longitud 4

sentido : Tipo de sentido de circulación de la calzada

Lista de Valores : 1 Sentido único (en la dirección de digitalización) 2 Doble sentido

Tipo de Valor Integer

Longitud 4

# 4. Vistas e índices.

Para mejorar la rapidez, hemos creado 2 vistas materializadas en base a la tabla punto, que es la más completa, y varios índices.

La consulta con la que se generan estas vistas es la siguiente:

-- CREATE EXTENSION unaccent;

-- View: public.puntos\_direccion

CREATE MATERIALIZED VIEW public.mv\_puntos\_direccion AS

SELECT punto.id,

st\_y(st\_transform(punto.punto, 4326)) AS latitud,

st\_x(st\_transform(punto.punto, 4326)) AS longitud,

st\_transform(punto.punto, 4326) AS p4326,

punto.idcalle,

calle.nombre,

punto.numero,

puntos\_localidades.nombre AS localidad,

puntos\_departamento.depa,

punto.padron AS ppadron,

solar.solar,

solar.padron AS spadron,

solar.manzana,

punto.km,

punto.nombre\_inmueble,

punto.cp,

to\_tsvector('spanish'::regconfig, (((((unaccent(COALESCE(calle.nombre, ' '::text)) || ' '::text) || unaccent(COALESCE(punto.nombre\_inmueble, ' '::text::character varying)::text)) || ' '::text) || unaccent(COALESCE(puntos\_localidades.nombre, ' '::text::character varying)::text)) || ' '::text) || unaccent(COALESCE(puntos\_departamento.depa, ''::text))) AS fulltext

FROM punto

LEFT JOIN calle ON punto.idcalle = calle.id

LEFT JOIN puntos\_localidades ON punto.id = puntos\_localidades.id

LEFT JOIN solar ON punto.solar\_id = solar.apis\_id

LEFT JOIN puntos\_departamento ON punto.id = puntos\_departamento.id;

-- limit 100;

CREATE INDEX idx\_fulltext\_mv\_puntos\_direccion

ON public.mv\_puntos\_direccion USING gin

(fulltext);

CREATE MATERIALIZED VIEW public.mv\_direcciones2

AS

SELECT DISTINCT count(\*) AS numregs,

mv\_puntos\_direccion.idcalle,

mv\_puntos\_direccion.nombre,

alias.buscar,

mv\_puntos\_direccion.localidad,

mv\_puntos\_direccion.depa AS departamento,

avg(mv\_puntos\_direccion.latitud) AS lat,

avg(mv\_puntos\_direccion.longitud) AS lng,

mv\_puntos\_direccion.nombre\_inmueble,

to\_tsvector('spanish'::regconfig, (((((unaccent(COALESCE(alias.buscar, ' '::text)) || ' '::text) || unaccent(COALESCE(mv\_puntos\_direccion.nombre\_inmueble, ' '::text::character varying)::text)) || ' '::text) || unaccent(COALESCE(mv\_puntos\_direccion.localidad, ' '::text::character varying)::text)) || ' '::text) || unaccent(COALESCE(mv\_puntos\_direccion.depa, ''::text))) AS fulltext

FROM mv\_puntos\_direccion

JOIN alias ON alias.idcalle = mv\_puntos\_direccion.idcalle

GROUP BY mv\_puntos\_direccion.nombre, alias.buscar, mv\_puntos\_direccion.idcalle, mv\_puntos\_direccion.localidad, mv\_puntos\_direccion.depa, mv\_puntos\_direccion.fulltext, mv\_puntos\_direccion.manzana, mv\_puntos\_direccion.nombre\_inmueble

HAVING mv\_puntos\_direccion.manzana IS NULL

WITH DATA;

CREATE INDEX idx\_fulltext\_mv\_direcciones2

ON public.mv\_direcciones2 USING gin

(fulltext)

TABLESPACE pg\_default;

CREATE INDEX idx\_puntos\_direccion\_p4326 ON public.mv\_puntos\_direccion USING GIST(p4326);

CREATE INDEX idx\_depa\_localidad\_nombre ON mv\_puntos\_direccion (depa, localidad, nombre);

CREATE INDEX mv\_direcciones2\_idcalle\_idx ON public.mv\_direcciones2 USING btree (idcalle);

La primera vista materializada (mv\_puntos\_direccion) se genera a partir de la tabla “punto” y sus enlaces (joins) con las tablas de “calle”, “puntos\_localidades”, “solar” y “puntos\_departamento”. Esta es la vista de la que obtendremos la ubicación final.

A partir de esta vista, generamos otra vista materializada: “mv\_direcciones2”, que es la que vamos a emplear para la búsqueda de candidatos. La vista la generamos con una consulta de tipo DISTINCT sobre los campos ya existentes de “mv\_puntos\_direccion”, y hacemos un JOIN con la tabla “alias” para buscar todas los variantes de una calle.

Esta tabla es la que usamos para hacer la búsqueda de candidatos. Una vez que el usuario selecciona una calle, entonces usamos el campo idcalle para obtener todos los números de portal de la vista “mv\_puntos\_direccion”, y buscar ahí el portal en concreto si el usuario lo ha especificado en la consulta.

El índice espacial (“idx\_puntos\_direccion\_p4326”) que creamos sirve para que la búsqueda de geocodificación inversa sea muy rápida.

# 5. Consideraciones Finales

Se deberían corregir (poco a poco) los problemas detectados en la base de datos y en los nuevos desarrollos, tener en cuenta los fallos y las optimizaciones del punto anterior.

Si se decide separar las direcciones, hay que eliminar el campo nombre\_inmueble de las vistas materializadas y en el servicio web, realizar las búsquedas de manera separada.

En cuanto a la tabla ruta, se aconseja seguir un modelo similar al que se usa en Cartociudad, y unificar ruta y calle en una única tabla “vial” (pág. 20 del documento de especificaciones de Cartociudad).

La idea principal es tener una única tabla “tramo” con campo geometría de tipo multilinestring. Los viales (calles o rutas) están formados por colecciones de tramos, que enlazan por el campo idcalle de tramo.

# 6. Scripts de Migración

En el GitLab se incluyen una serie de Scripts SQL para generar la el nuevo esquema “budu” con las tablas, campos, procedimientos almacenados, funciones, extensiones, triggers, etc.

Para aplicar los cambios, seguir las instrucciones que hay en el repositorio (archivo Leeme.txt del directorio ocp/SQL.