MÁSTER UNIVERSITARIO EN SOFTWARE LIBRE

Rendimiento Web Gis en relación a Servidores de Mapas y sus orígenes de datos

TRABAJO FINAL DE MASTER

Autor: Gustavo Navas

Consultor UOC: Antoni Pérez Navarro

Junio 2014

Contenido protegido por una licencia Creative Commons BY-SA 4.0, que permite su copia y difusión ilimitadas, y uso en obras propias con la misma licencia CC. La licencia completa puede verse en:

http://creativecommons.org/licenses/by-sa/4.0

Agradecimientos:

Quiero agradecer a mi consultor Antoni Pérez, que siempre fue certero, preciso en sus observaciones, comentarios y sugerencias.

A los miembros del grupo de investigación IDE-UPS, sobre Infraestructura de datos y Geoportales del CIMA, de la Universidad Politécnica Salesiana, en especial a Rodrigo y Edwin.

A mi esposa, que sin su apoyo y amor, no hubiera podido terminar este trabajo.

Dedicatoria:

A mi familia, mi esposa Elenita, el amor de mi vida, a mis hijos Gustavo Sebastián y Gabriel Ernesto.

A la memoria de mi hermana Isabel, su fuerza de vida siempre es una motivación en mi espíritu y mi corazón.

RESUMEN:

El presente trabajo tiene el propósito de estudiar las herramientas para geoportales en relación a su rendimiento. Para esto se consideran los siguientes parámetos: servidor de mapas como Geoserver y Mapserver y orígenes de datos como son Shapefile, PostGIS, GML, GeoJson, y WMS. Llevando esto a establecer como las partes gráfica y textual son manejadas en los diferentes servicios, también es importante destacar el manejo de la información geográfica a nivel del cliente utilizando OpenLayer.

Palabras Claves:

WebGIS, Geoportales, Rendimiento, Servidores de Mapas, PostGIS, ShapeFile, GML, GeoJSon.

Índice de Contenido

1. INTRODUCCION 1
2. INVESTIGACIONES PREVIAS 2
3. FUNDAMENTOS TEORICOS 2
3.1 MANEJO DE MAPAS EN ENTORNOS WEB 3
3.2 ORIGENES DE DATOS 3
3.3 RENDIMIENTO 3
3.3.1 Herramientas empleadas 4
4. METODOLOGIA Y PROCEDIMIENTO 4
4.1 PROCESO. 5
4.1.1 Instalación del servidor y herramientas a ser empleadas. 5
4.1.2 Instalación de herramientas en cliente. 5
4.1.3 Selección de datos geográficos y subida al servidor 5
4.1.4 Modificación de archivos 6
4.1.5 Arquitectura 6
4.1.6 Pre prueba: Empleando Firebug, determinar los tamaños de datos y velocidad de respuesta. 7
4.1.7 Pruebas JMeter 8
5 ANALISIS DE RESULTADOS 11
5.1 OpenLayer 11
5.2 WMS 12
5.3 GML: PostGIS vs. ShapeFile 13
5.4 GML y GeoJson 14
5.5 Mapserver, Geoserver y OpenLayer 15
5.6 Mapserver: ShapeFile y PostGIS 16
5.7 Decimales 16
6. DISCUSIÓN Y CONCLUSIONES 18
7. REFERENCIAS 20
ANEXO 1 23

Índice de Figuras

Figura 1 Arquitectura utilizada para Geoserver y Mapserver 7
Figura 2 Ejemplo de datos proporcionados por Firebug. 8
Figura 3 Pantalla inicial de JMeter donde se configuran los parámetros, los resultados obtenidos y el gráfico 9
Figura 4 Tamaño de archivo (parte gráfica) vs rendimiento en milisegundos 11
Figura 5 Tamaño de archivo (parte gráfica) vs rendimiento y tamaño de archivo (parte gráfica + textual) vs rendimiento 12
Figura 6 Rendimiento OpenLayer y WMS par origenes de datos Shapefile y PostGIS 13
Figura 7 Relación de rendimientos PostGIS vs ShapeFile 14
Figura 8 Relación de rendimientos GeoJson vs GML 15
Figura 9 Comparación entre tamaños vs. rendimiento para diferentes origenes de datos 15
Figura 10 Comparación de rendimientos en Mapserver empleando JMeter y Firebug 16
Figura 11 Manejo de decimales 17
Figura 12 Cruce entre archivos donde el uno se manda solo el gráfico y el otro solo el texto, realizado con datos del DMQ 18

Índice de Tablas

Tabla 1 Diferentes orígenes de datos en formato ShapeFile 10
Tabla 2 Tamaño Archivo vs. rendimiento 11
Tabla 3 Para una entrega del servidor Geoserver en formato GML empleando orígenes de datos correspondientes a DMQ y con los datos originales en ShapeFile y PostGIS 13
Tabla 4 Rendimiento GeoJson y GML 14
Tabla 5 Decimales 17

1. INTRODUCCION

Los geoportales o Web GIS como se los llama en inglés, son la forma de mostrar cualquier recurso resultado de alguna investigación, o parámetro que se desee poner sobre una ubicación geográfica.

El geoportal se ha convertido en el portal Web, que se adapta a muy diversos campos del conocimiento y la investigación. Poder tener la oportunidad de ubicar en un entorno de internet el resultado de un Sistema de Información donde uno de los parámetros es su ubicación geográfica es cada vez más importante, es por tanto necesario el estudio de los geoportales en forma sistemática.

Los servidores de mapas y los orígenes de datos, son analizados en el presente trabajo en función de su rendimiento, se estudiará el Mapserver, Geoserver y el OpenLayer. Además se analiza el hecho de que los datos geográficos están formados por dos partes, una gráfica y otra textual y su influencia en el rendimiento de un Geoportal.

En Ecuador, uno de los mayores retos como país, consiste en disponer de información de la mejor calidad y actualizada, para ello existe el SNI, Sistema Nacional de Información, que en su parte fundamental reza "constituye el conjunto organizado de elementos que permiten la interacción de actores con el objeto de acceder, recoger, almacenar y transformar datos en información relevante para la planificación del desarrollo y las finanzas públicas" Del Código de Planificación y Finanzas Públicas, Art.33 Registro Oficial No.306, año 2010 (Sistema Nacional de Información, 2014).

Dos fuentes de información son fundamentales para esta institución, la información estadística y la información Geográfica. Adicionalmente muchas instituciones en el Ecuador, tienen su Geoportal, y en especial los GAD's "Gobiernos Autónomos Descentralizados", como son alcaldías y prefecturas. El presente estudio puede ser una guía útil para escoger las herramientas y también analizar los datos geográficos a ser puestos en un geoportal.

Este artículo presenta la siguiente estructura: en el numeral 2 se hace una revisión de las investigaciones con relación al rendimiento de geoportales; el 3 trata los fundamentos teóricos requeridos para la presente investigación; el 4 se revisa la metodología y procedimiento utilizado en la investigación; el 5 presenta los resultados y el 6 contiene las conclusiones.

2. INVESTIGACIONES PREVIAS

El rendimiento en geoportales tiene total vigencia. Al momento existe una gran cantidad de investigaciones que tratan este tema, pero en la mayoría de ellos se lo hace para resolver un problema específico y todos los esfuerzos se encaminan en esa línea, no se aprecian trabajos donde se analicen en forma general las herramientas disponibles como son el caso de los servidores. Una excepción a este tipo de investigación es el trabajo cuyo título es "Evaluación del rendimiento de los servicios WMS de Mapserver y Geoserver para la implementación IDE" (Naranjo Martinez, 2014), este es una investigación sobre el rendimiento con 4 escenarios posibles, en todos ellos la variable que se considera para el eje X es el número de usuarios concurrentes a un servicio WMS tanto en Geoserver y Mapserver

El presente trabajo analiza diversos orígenes de datos como son: Shapefile, PostGIS, GML, GeoJson, servicios WMS, etc. A continuación se mencionan investigaciones que utilizan estos o alguno de estos orígenes de datos, como por ejemplo el artículo de (Hong & Liao, 2011) que trabaja con información en GML, KLM, XML y GeoJson, en otro artículo se menciona a PostGIS y GeoJson, como lenguaje script PHP, y la plataforma para manejo de mapas Open Street Map (Piccinini, Smith, Hooke, & Hesketh, 2013).

En el artículo sobre la herramienta Geopot (Lee & Liang, 2011), se estudia la "Evaluación de Ejecución" y se realizan una serie pruebas de rendimiento entre PostGIS vs. GeoPot.

Otra investigación que sirvió de guía para el presente trabajo es el artículo sobre transmisión de datos vectoriales a través del internet. (Yang, Purves, & Weibel, 2007), los pasos que ellos siguen son:

1 Clasificación de los vértices para la eliminación potencial
2 Reglas para la comprobación de la validez de topología
2.1 Operaciones para la eliminación de vértices.
3 Estructura de datos para la transmisión y la reconstrucción
4 Arquitectura de Aplicación
5. Resultados y discusión
5.1 experimentos
5.2 Rendimiento de la eficiencia de transmisión (Yang, Purves, & Weibel, 2007)

3. FUNDAMENTOS TEORICOS

A continuación algunos conceptos que se usan dentro de esta investigación.

3.1 MANEJO DE MAPAS EN ENTORNOS WEB

Los dos servidores de mapas más utilizados en la actualidad son el MapServer y el GeoServer. El MapServer es "una plataforma de Código Abierto para la publicación de datos espaciales y aplicaciones cartográficas interactivas para la web. Es publicado bajo licencia tipo MIT y funciona en los principales sistemas operativos (Windows, Linux, Mac OS X)" (MapServer, 2014).

Con respecto a GeoServer este es "un servidor de código abierto escrito en Java, permite a los usuarios compartir y editar datos geospaciales. Diseñado para la interoperabilidad, publica datos de las principales fuentes de datos espaciales usando estándares abiertos" (Wikipedia, GeoServer, 2013).

Por otro lado existe OpenLayer que es una biblioteca en JavaScript para mostrar mapas interactivos en un portal de internet (OpenLayers, 2013); este se carga en el navegador, logrando que muchas acciones las realice el cliente, evitando de esa manera la congestión en el servidor por tener que procesar todos los pedidos.

La calidad de los datos, su manera de almacenarse, registrarse y mostrarse está controlado a través de la norma ISO19113 del año 2002 (Hong & Liao, 2011)

3.2 ORIGENES DE DATOS

Los datos geográficos que se utilizan en entornos GIS son muy variados, sin embargo el origen inicial de estos datos comúnmente son los Shapefiles, también están disponibles en la base de datos PostGIS versión 2.01 (PostGis, 2013), procesados como un servicio WMS o provistos en formatos vectoriales especializados para entornos Web como GeoJson, GML, entre otros.

Independientemente de ello, la información geográfica siempre tiene dos elementos fundamentales: una parte gráfica y una parte textual, que son claramente identificables en el formato Shapefile.

3.3 RENDIMIENTO

El tiempo total desde que el cliente hace una solicitud al servidor, hasta que recibe la información y esta se muestra en el navegador, se compone por el tiempo que toma cada componente en realizar su trabajo específico. Por lo tanto el tiempo puede ser descrito como la sumatoria de cada una de las actividades que desarrolla el servidor y el cliente hasta completar y mostrarse el resultado en el navegador. Esta información se encuentra

en el capítulo sobre rendimiento del libro de la serie" ISPRS Book Series" (Lee & Liang, 2011), cuyo título es "Advanced in Web-based GIS, Mapping Services and Application".

En dicho libro el tiempo total se determina a través de la siguiente ecuación:
.

$$T = T_{envio} + T_{proceso} + T_{recepción}$$

Ecuación 1

Todas estas actividades podrían resumirse en tres:

1. Envío de la solicitud al servidor.
2. Procesamiento de la solicitud en el servidor.
3. Entrega de información al cliente.

3.3.1 Herramientas empleadas

Dentro del proceso se emplearon la herramienta Firebug para las prepruebas y Apachache JMeter para las pruebas:

Firebug es "una extensión de Firefox creada y diseñada especialmente para desarrolladores y programadores web. Es un paquete de utilidades con el que se puede analizar (revisar velocidad de carga, estructura DOM), editar, monitorizar y depurar el código fuente, CSS, HTML y JavaScript de una página web". (Wikipedia, 2014)

El JMeter "es una herramienta de prueba de carga para analizar y medir el desempeño de una variedad de servicios, con énfasis en aplicaciones web". (Wikipedia, 2013).

4. METODOLOGIA Y PROCEDIMIENTO

La investigación realizada ha sido de tipo experimental, con las siguientes características:

Paradigma: Positivista

"El postulado fundamental del positivismo es que el conocimiento válido sólo puede establecerse por referencia a lo que se ha manifestado a través de la experiencia." (Centro Virtual Cervantes, 2014).

Diseño: Experimental

Se enmarca en el diseño experimental porque en primer lugar se plantearon las hipótesis a ser probadas, luego de esto se manipularon las variables para medir el

efecto que esto causaba en los resultados observados y a través del manejo de cuadros estadísticos se obtuvieron los resultados. (Oates, 2008)

Metodología: Cuantitativa

"Datos cuantitativos significan datos o evidencia basado en números, este es el principal tipo de datos generados por experimentos y encuesta" (Oates, 2008)

Análisis: Estadístico

"La Estadística estudia los métodos científicos para recoger, organizar, resumir y analizar datos, así como para sacar conclusiones válidas y tomar decisiones razonables basadas en el análisis", (Murray, 2009).

4.1 PROCESO.

A continuación se describe el proceso seguido:

4.1.1 Instalación del servidor y herramientas a ser empleadas.

a. Servidor Centos 6
    i. Nombre: CentosIDE03
    ii. ip privada: 172.17.128.196
    iii. ip pública: 190.15.136.4
    iv. http://190.15.136.4
b. Apache 2.0
c. Apache Tomcat 7.0.52
d. Postgresql 9.3.4 – Postgis 2.1.2
e. Mapserver 5.6.9
    http://190.15.136.4:8080
f. Geoserver 2.3.5
    http://190.15.136.4:8081/geoserver

4.1.2 Instalación de herramientas en cliente.

a. Computador: Laptop Dell Inspiron 5110 i5, con sistema operativo Windows 7
b. Firebug 1.12.18 sobre Mozilla Firefox 29.0.
c. Apache JMeter 2.11

4.1.3 Selección de datos geográficos y subida al servidor

a. Los datos geográficos recopilados se los puede observar en la Tabla 1, Estos fueron subidos al servidor en dos formatos el uno en su original ShapeFile y el otro a la base de datos PostGIS. A estos dos orígenes de datos los servidores

Geoserver y/o Mapserver acceden para procesarlos y entregar la información geográfica al cliente o navegador.

b. Geoserver entrega al navegador en tres formatos posibles dos vectoriales que son GML, y GeoJson y un servicio WMS.
c. Además se utilizó el OpenLayer consumiendo servicios WMS.
d. Para Mapserver este entrega una imagen a partir de acceder tanto al ShapeFile como a la base de datos Postgis.

4.1.4 Modificación de archivos

a. Como la mayoría de análisis a realizarse relacionan el tamaño de los archivos con su rendimiento, se tuvo la necesidad de modificar datos de información geográfica para obtener valores intermedios, manteniendo siempre su estructura original, ya que se necesitaba disponer de una secuencia numérica para los valores de tamaños de archivo y poder ponerlos sobre el eje X, de las gráficas analizadas.
b. Los archivos de dominio público tenían parámetros distintos es por esto que resultaba difícil hacer comparaciones entre ellos, por ejemplo tenían
    i. Diferentes sistemas de coordenadas.
    ii. Diferentes tipos de datos: punto, línea o polígono.
    iii. Información dispersa.
c. En el análisis de la parte gráfica y textual los archivos creados fueron los que mejor resultado dieron debido a que se podía tener un mejor control de la relación de tamaño de la parte gráfica vs la textual.
d. Estos archivos modificados también fueron subidos al servidor y tuvieron el mismo tratamiento explicado en el apartado 3.
e. El anexo 1 muestra las modificaciones realizadas en dos orígenes de datos geográficos, el uno de nombre pobreza al cual se lo modificó reduciendo el número de provincias y el otro del Distrito Metropolitano de Quito, al cual utilizando las herramientas de simplificación y densificación de Quantum GIS se generaron los datos geográficos correspondientes.

4.1.5 Arquitectura

La Figura 1 muestra la arquitectura utilizada tanto para Mapserver como para Geoserver.

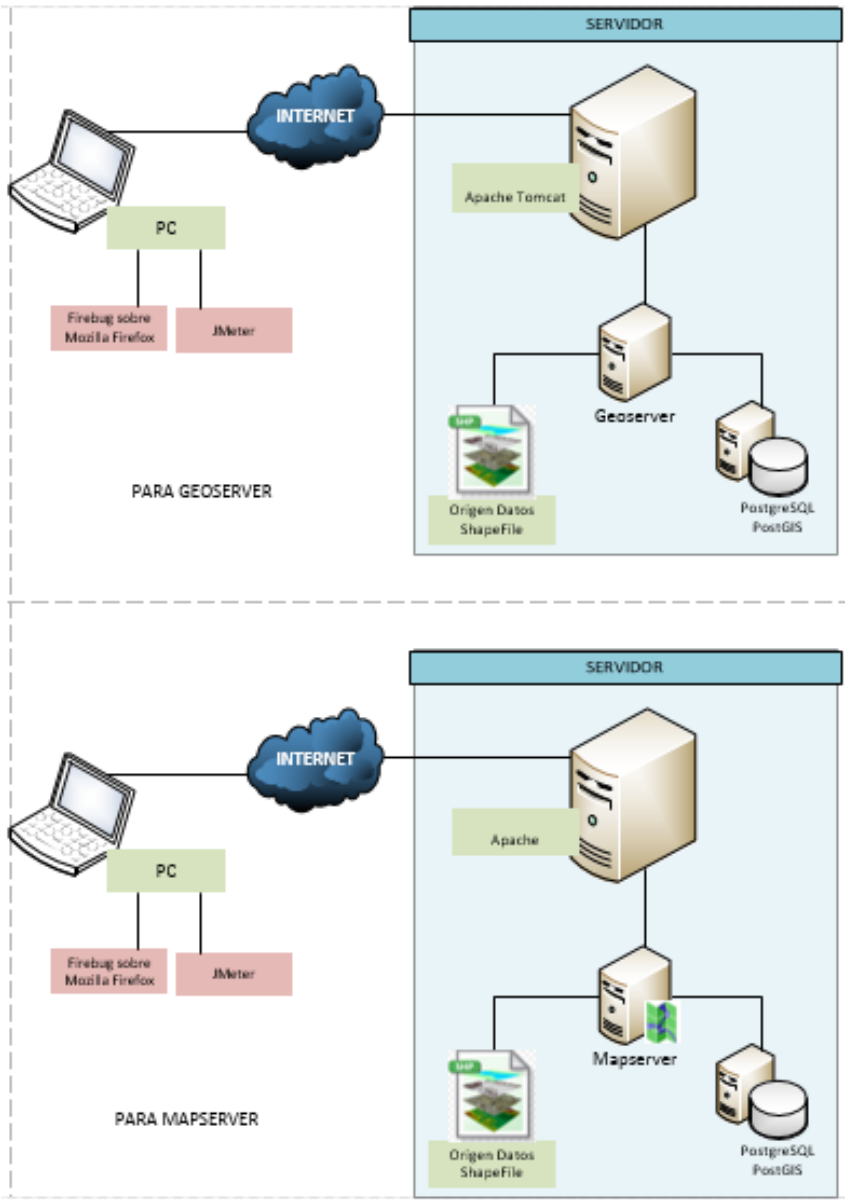


Figura 1 Arquitectura utilizada para Geoserver y Mapserver

4.1.6 Pre prueba: Empleando Firebug, determinar los tamaños de datos y velocidad de respuesta.

a. La herramienta Firebug que corre dentro de Mozilla Firefox, muestra en detalle cada uno de los tiempos de los procesos, un ejemplo de lo dicho se puede apreciar en la Figura 2. Estas pruebas se las llevó a cabo sin ejecutar ningún otro proceso en el computador.

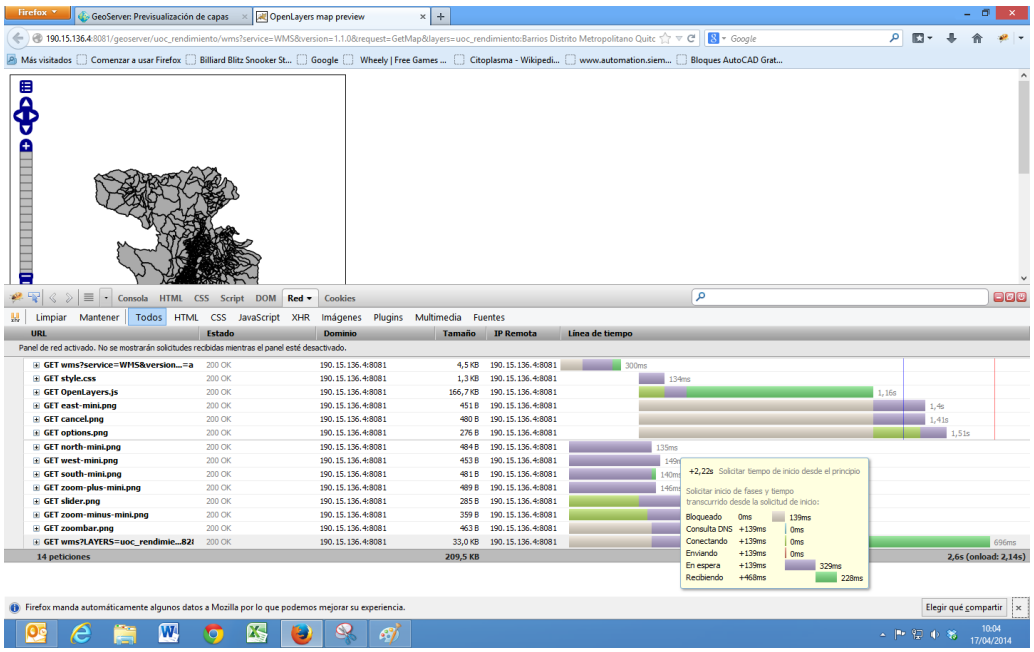


Figura 2 Ejemplo de datos proporcionados por Firebug.

4.1.7 Pruebas JMeter

a. En JMeter se establecieron todas las pruebas que se hicieron luego de las prepruebas con Firebug. El Firebug dio el valor de tamaño y la velocidad de respuesta de una sola corrida mientras que con el JMeter se tuvo la información con un número de 300 muestras.
b. Con cada una de las pre pruebas realizadas en Firebug, se realizó una corrida en JMeter.
   Con las direcciones URL se configuró el JMeter manteniendo las condiciones que se describen a continuación tanto en el cliente como en el servidor, se establecieron los siguientes controles para la realización de las pruebas:

   1. Hardware: el servidor, así como también el cliente siempre fueron los mismos, para garantizar que la experimentación no se afecte por este motivo.
   2. Al estudiar un servidor de mapas, se mantuvo el otro servidor de mapas inactivo, para que no influya el uno sobre el otro.
   3. El cliente solo ejecutó las herramientas con la que se realizaron las pruebas.
   4. Con relación a la comunicación, esta se llevó a cabo en una misma red.
   5. Antes de las pruebas, se instalaron todas y cada una de las aplicaciones y elementos que forman parte de las pruebas, esto tanto para el servidor, como para el cliente.

Tanto para los archivos correspondientes al DMQ así como para los archivos de la situación Socio Económica en el Ecuador se mantuvo en JMeter el número de muestras constante en 300, ver Figura 3, y esta información fue recopilada en una hoja electrónica para poder realizar el análisis de los resultados.

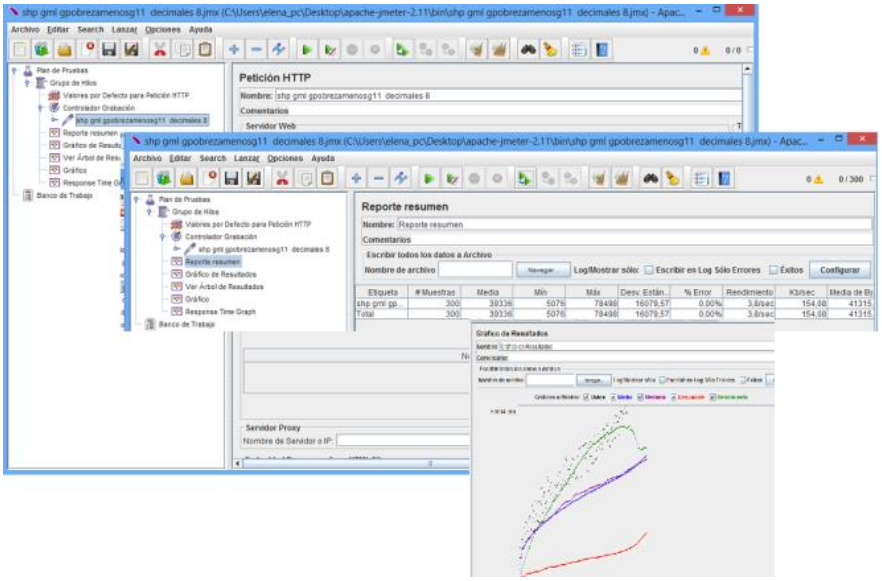


Figura 3 Pantalla inicial de JMeter donde se configuran los parámetros, los resultados obtenidos y el gráfico.

Tabla 1 Diferentes orígenes de datos en formato ShapeFile

	aeropuerto	pobreza	DMQ parroquias	ferrocarril	islas	lago-laguna	límite-costanero	límite	Poblados	represas	ríoDoble	río_ torrente	vías
Formato Shape	Puntos	Polígono	Polígono	Líneas	Polígono	Polígono	Línea	Línea	Puntos	Polígono	Polígono	Líneas	Líneas
No. Atributos	5	1	4	4	3	5	3	2	5	3	5	5	3
Número Filas	105	1	1269	4	529	257	1	1	2675	6	40	2093	2376
archivo shp (KB)	3	4354	3368	18	1207	66	59	797	74	32	157	2802	353
archivo dbf(KB)	106	1142	81	3	395	258	1	1	3321	5	41	2098	1771
archivo shx (KB)	1	8	11	1	5	3	1	1	21	1	1	17	19
archivo prj (KB)	1		1	1	1	1	1	1	1	1	1	1	1
idx (KB)													
sbn (KB)		9	13										
sbx (KB)		1	1										
shp.xml (KB)			124										
TOTAL (KB)	111	5514	3599	23	1608	328	62	800	3417	39	200	4918	2144
Porcentaje shp	2,70%	78,96%	93,58%	78,26%	75,06%	20,12%	95,16%	99,6%	2,17%	82,05%	78,50%	56,97%	16,5%
Porcentaje dbf	95,50%	20,71%	2,25%	13,04%	24,56%	78,66%	1,61%	0,1%	97,19%	12,82%	20,50%	42,66%	82,6%
TOTAL shp y dbf	109	5496	3449	21	1602	324	60	798	3395	37	198	4900	2124
Porcentaje shp	2,8%	79,22%	97,7%	85,7%	75,3%	20,4%	98,3%	99,9%	2,2%	86,5%	79,3%	57,2%	16,6%
Porcentaje dbf	97,2%	20,78%	2,3%	14,3%	24,7%	79,6%	1,7%	0,1%	97,8%	13,5%	20,7%	42,8%	83,4%
Tamaño Texto . Tamaño Gráfico	**35,33**	**0,26**	**0,02**	**0,17**	**0,33**	**3,91**	**0,02**	**0,00**	**44,88**	**0,16**	**0,26**	**0,75**	**5,02**

5 ANALISIS DE RESULTADOS

5.1 OpenLayer

Al usar la combinación del servicio WMS del Geoserver junto con el OpenLayer, y un número de 300 muestras corridas en JMeter el rendimiento tiende a estabilizarse en un valor alrededor de 29.1 (Peticiones/seg), o 306.62 milisegundos. Sin importar el tamaño del origen de datos el rendimiento para las 300 muestras resultó ser la mejor opción y el valor se mantuvo relativamente estable y bajo, la Tabla 2 muestra los valores extremos de tamaño de dos shapefiles. La Figura 4, confirma lo dicho en el párrafo anterior pequeña variación del rendimiento frente al tamaño.

Tabla 2 Tamaño Archivo vs. rendimiento

Tamaño (KB)	Rendimiento (ms)
18	290
26252	312

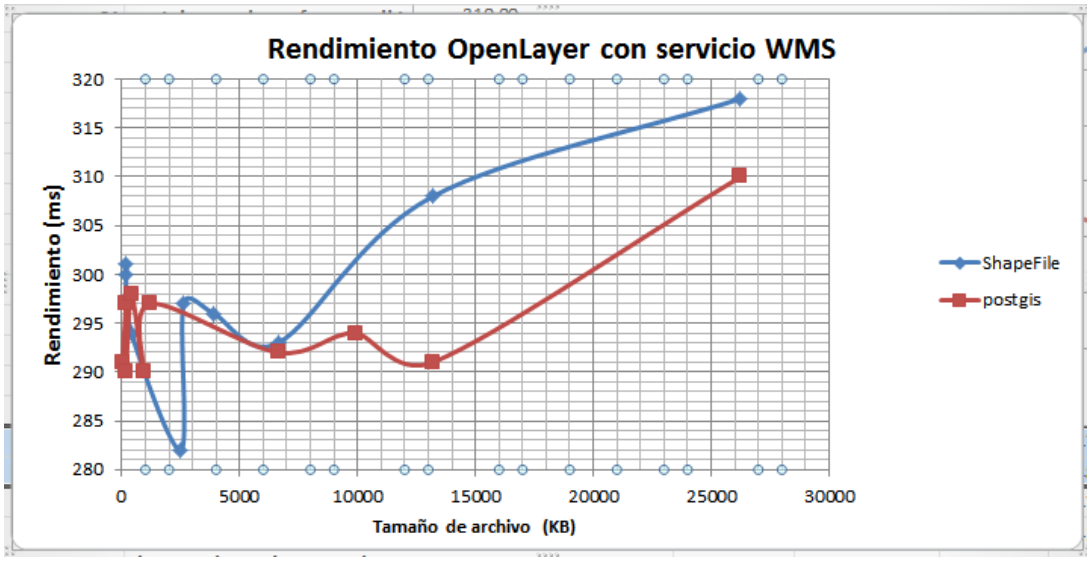


Figura 4 Tamaño de archivo (parte gráfica) vs rendimiento en milisegundos

El dato para el tamaño de archivo que se usó en la Figura 4, corresponde solo a su parte gráfica, esto debido a que se determinó que para el Servicio WMS, la parte textual no tiene ningún efecto sobre el rendimiento, esto se puede comprobar en la Figura 5, que tiene las

mismas curvas de la Figura 4, más dos curvas adicionales en la que como tamaño se da la suma de los dos archivos (shp + dbf)..

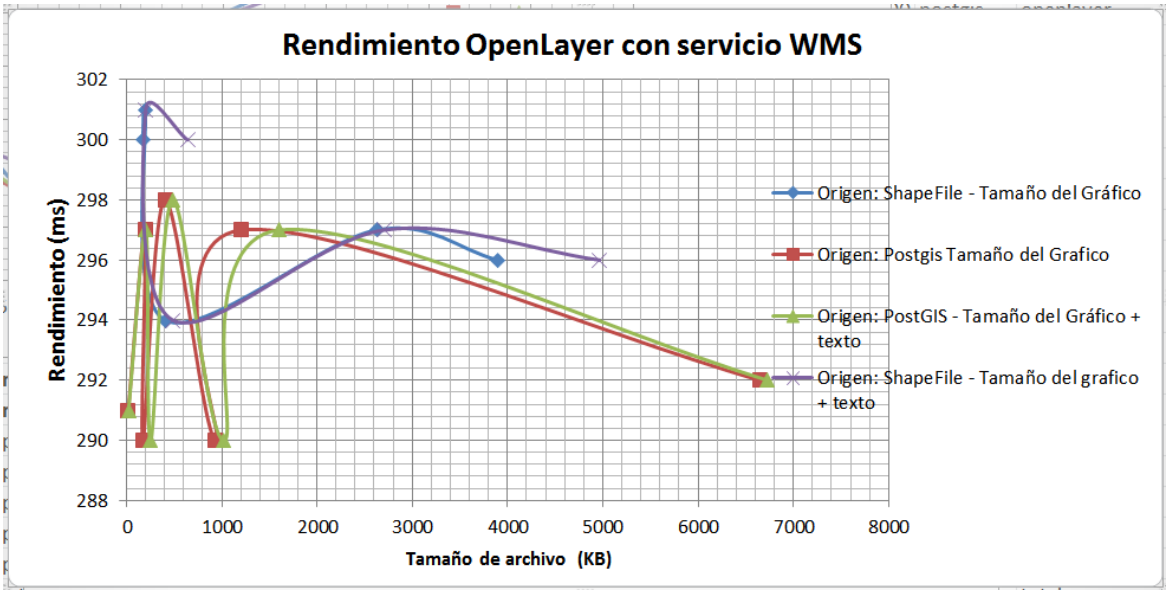


Figura 5 Tamaño de archivo (parte gráfica) vs rendimiento y tamaño de archivo (parte gráfica + textual) vs rendimiento

5.2 WMS

En servicios WMS cada llamada es un servicio atendido a nivel de servidor. Cuando el origen de datos es pequeño este tiende a ser constante para el rendimiento, pero si el origen de datos es grande el rendimiento se ve considerablemente disminuido, esto se puede apreciar en la Figura 6, los datos corresponden a aeropuerto con un tamaño de 109 KB y vegetación que corresponde a 29461 KB.

Con OpenLayer ejecutándose, también el rendimiento tiende a ser constante en un valor máximo debido a que la mayoría se ejecutará a nivel de cliente y se entiende que el cliente conserva la imagen entregada por el servicio WMS. Esto se puede ver en la Figura 6.

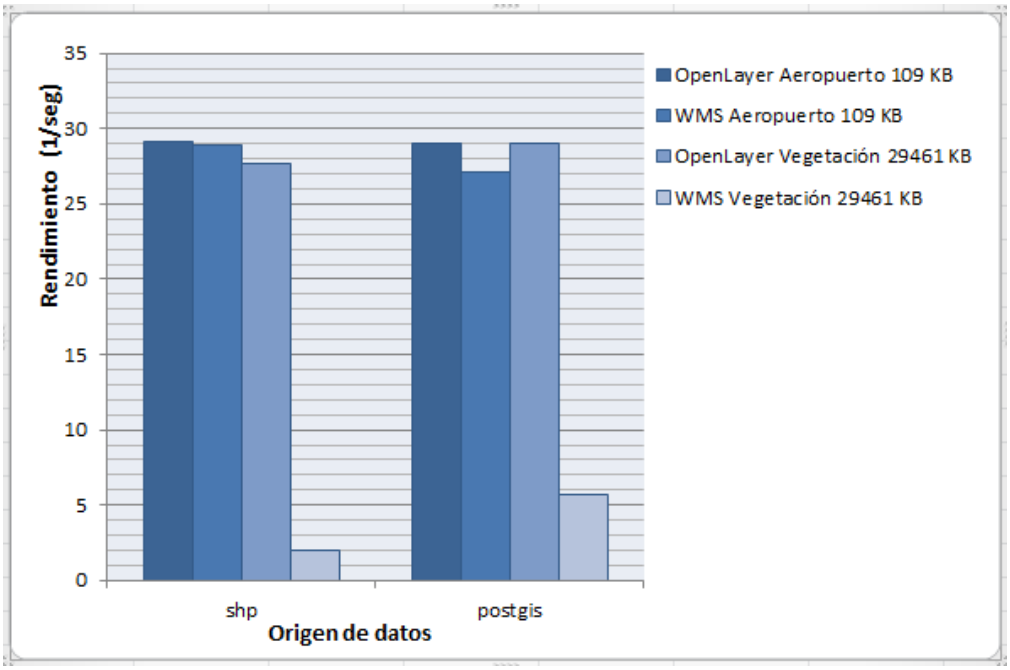


Figura 6 Rendimiento OpenLayer y WMS par origenes de datos Shapefile y PostGIS

5.3 GML: PostGIS vs. ShapeFile

Se estableció una serie de pruebas para encontrar relaciones entre los rendimientos para diversos orígenes de datos. Al probar esto sobre Geoserver se logró establecer que para un origen de datos como los que se muestran en la Tabla 3, el rendimiento medido en peticiones/seg, para un origen inicial de ShapeFile vs. un origen PostGIS, la relación es lineal leyendo estas tenían una relación lineal como muestra la Figura 7

Tabla 3 Para una entrega del servidor Geoserver en formato GML empleando orígenes de datos correspondientes a DMQ y con los datos originales en ShapeFile y PostGIS

	Shp file	PostGIS
s.0001	25,90	28,90
DMQ	7,30	7,40
DMQ1	5,00	5,00
DMQ2	2,50	2,50
DMQ3	0,81	1,30

Es claro observar que los rendimientos coinciden en gran medida, para confirmarlo se ha graficado esta tabla como se lo puede apreciar en la Figura 7.

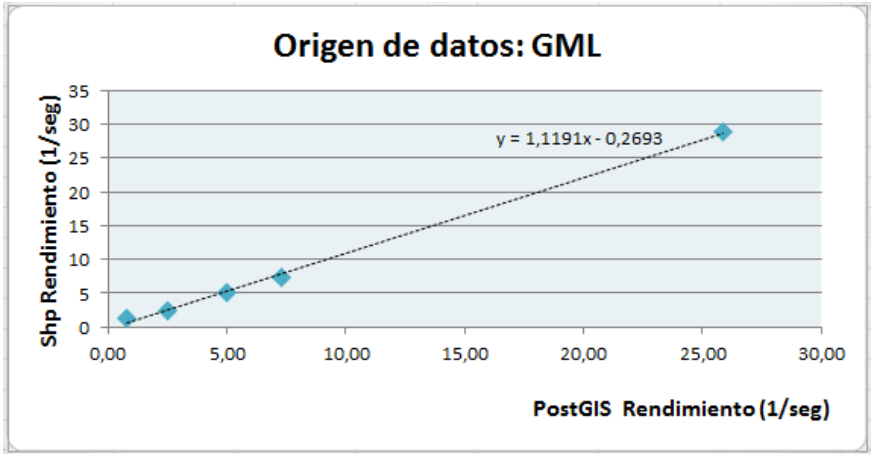


Figura 7 Relación de rendimientos PostGIS vs ShapeFile

5.4 GML y GeoJson

Trabajando con los datos del Distrito Metropolitano de Quito y sus diferentes formatos modificados, como muestra la Tabla 4, se observa los rendimientos para GeoJson y GML, graficando dicha información se puede observar que la relación ente ellos es de tipo exponencial hasta cierto valor y lineal a partir de allí, ver Figura 8 en donde se muestran las dos curvas y las ecuaciones que las gobiernan.

Tabla 4 Rendimiento GeoJson y GML

Geoserver	GeoJson rendimiento 1/segundo	GML rendimiento 1/segundo
s.01	5.50	29.10
s.0005	5.80	28.70
s.0001	2.90	29.00
Dmq1s.0001	2.62	23.83
Dmq2s.0001	1.40	10.34
DMQ	0.68	5.30
DMQ1	0.53	5.00
DMQ2	0.43	4.00

Para la relación exponencial el error porcentual resultó ser del 4.53%, que se considera aceptable, y en la relación lineal el error corresponde a un 0,47% que resulta ser muy satisfactorio. Esta consideración de “aceptable” y “muy satisfactorio” se lo tomó del documento “El proceso de medición: Análisis y comunicación de datos experimentales” (Santo & Lecumberry, 2005) .

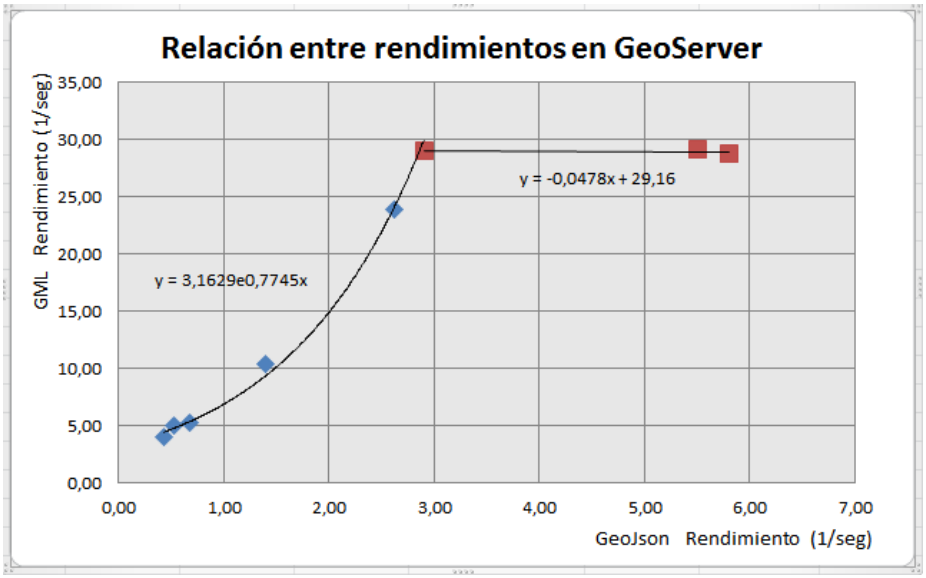


Figura 8 Relación de rendimientos GeoJson vs GML

5.5 Mapserver, Geoserver y OpenLayer

Como ya se dijo en un apartado anterior, el uso de OpenLayer con un servicio WMS, provisto por GeoServer, es el que muestra el mejor rendimiento, esto también se comprobó con los datos de "pobreza" y este valor se volvió constante alrededor de 19.5 (peticiones/seg) ver Figura 9 en esa misma grafica se confirmó que el Mapserver tiene una mejor velocidad de respuesta frente a llamadas al servicio WMS de Geoserver

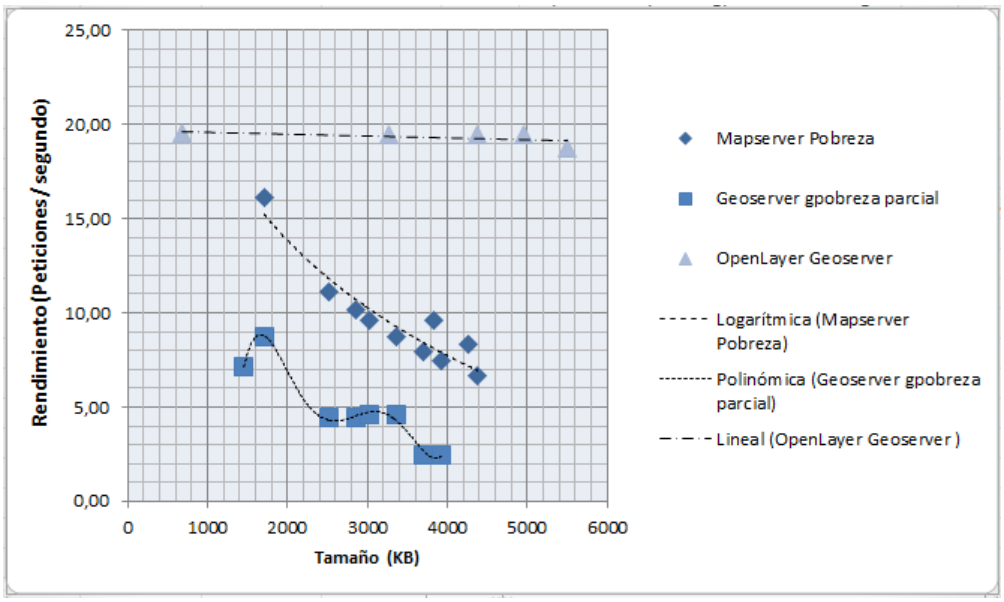


Figura 9 Comparación entre tamaños vs. rendimiento para diferentes origenes de datos

5.6 Mapserver: ShapeFile y PostGIS

Mapserver también maneja un valor de estabilidad cuando a partir de los ShapeFile se tiene que el rendimiento permanece con un valor casi constante alrededor de 285 ms, PostGIS, presenta mayor inestabilidad en función del tamaño de los orígenes de datos, esto se lo puede apreciar en la Figura 10, en esa misma gráfica también se pueden apreciar los valores tomados el Firebug, que correspondería a una sola llamada al servidor, esto tanto para ShapeFile como PostGIS.

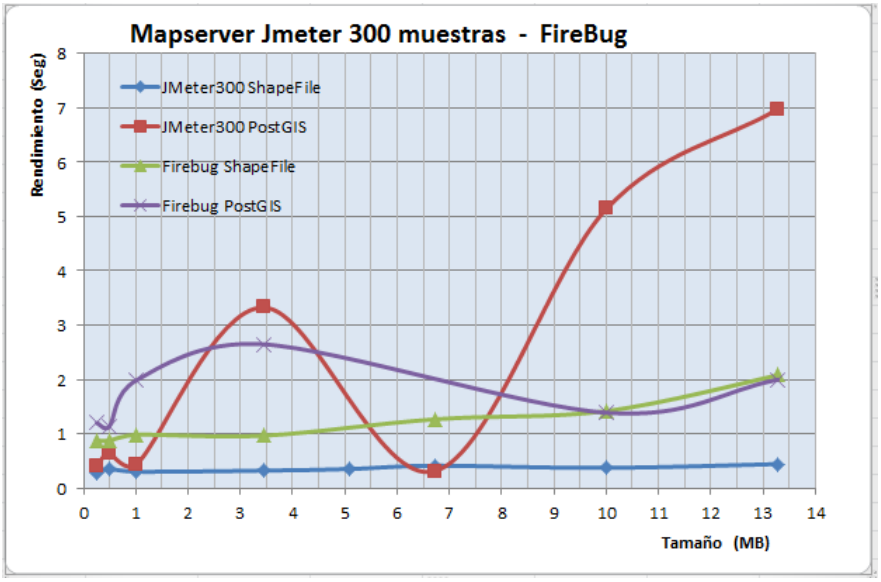


Figura 10 Comparación de rendimientos en Mapserver empleando JMeter y Firebug

Mapserver también maneja un valor de estabilidad cuando a partir de los ShapeFile se tiene que el rendimiento permanece con un valor casi constante alrededor de 285 ms, PostGIS, presenta mayor inestabilidad en función del tamaño de los orígenes de datos

5.7 Decimales

Al generar archivos vectoriales tipo GML y GeoJson estos tenían 8 decimales. Para las coordenadas del Distrito Metropolitano de Quito, que se encuentra en coordenadas geográficas, se puede decir que 0.1 segundos corresponden a un número 0.00003 ($3\ x\ 10^5$), pero para los archivos sobre pobreza estas se encuentran en metros (UTM) y 8 decimales resultan ser excesivos.

Se realizó con los datos de pobreza ("Situación Socio Económica del Ecuador"), una prueba en Geoserver, con datos tipo GML, se pudo ver que el número de decimales alteraba el tamaño del archivo y por tanto su rendimiento. En la Tabla 5 se aprecia el número de decimales empleados y su relación con el tamaño del archivo. Para mayor claridad se lo ha graficado, ver Figura 11

Tabla 5 Decimales

Origen Datos: GML		Servidor: GeoServer	
Número de decimales	**Gráfico y texto (KB)**	**Gráfico (KB)**	**Porcentaje Texto sobre total**
Cero	21,4	18,4	46,2%
Un	25,8	22,6	46,7%
Dos	30,2	26,8	47,0%
Tres	33,9	30,3	47,2%
Cuatro	37,5	33,4	47,1%
Cinco	41,0	36,5	47,1%
Seis	44,0	39,8	47,5%
Siete	46,9	43,1	47,9%
Ocho	49,8	46,3	48,2%

En esta Figura 11 es evidente como aumenta el tiempo de respuestas solo por el número de decimales empleados.

.

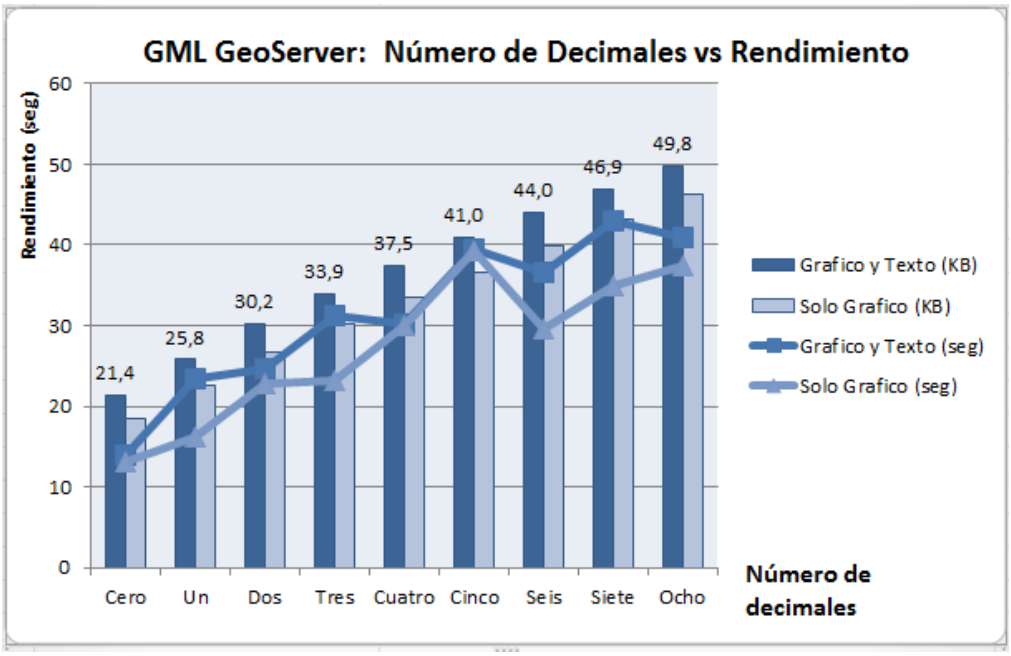


Figura 11 Manejo de decimales

6. DISCUSIÓN Y CONCLUSIONES

Primera Hipótesis

Enunciado: El momento de enviar una información geográfica a través de la web, existe un valor de relación entre el tamaño de información textual y el tamaño de información gráfica en el cual se tiene una considerable mejora en el rendimiento si solo se envía información gráfica al cliente desde el servidor. Es decir para una razón de tamaños de información textual e información gráfica mayor a cierto valor es mejor enviar solamente la información gráfica y para valores de la razón menores se enviará toda la información tanto gráfica como textual.

CONCLUSIÖN:

Para servicios WMS el cual procesa solamente la información gráfica este enunciado no tendría ningún efecto, pero para información geográfica de tipo vectorial enviados a través de la red como son GML y GeoJson si se puede encontrar el valor en el cual este tiene efecto. Es importante también mencionar que antes de enviar la información deberá analizarse el número de decimales con los que se debe enviar la coordenada, como ejemplo mencionemos que el ShapeFile de nombre pobreza está en un sistema de coordemanas UTM y la información inicial maneja 8 decimales.

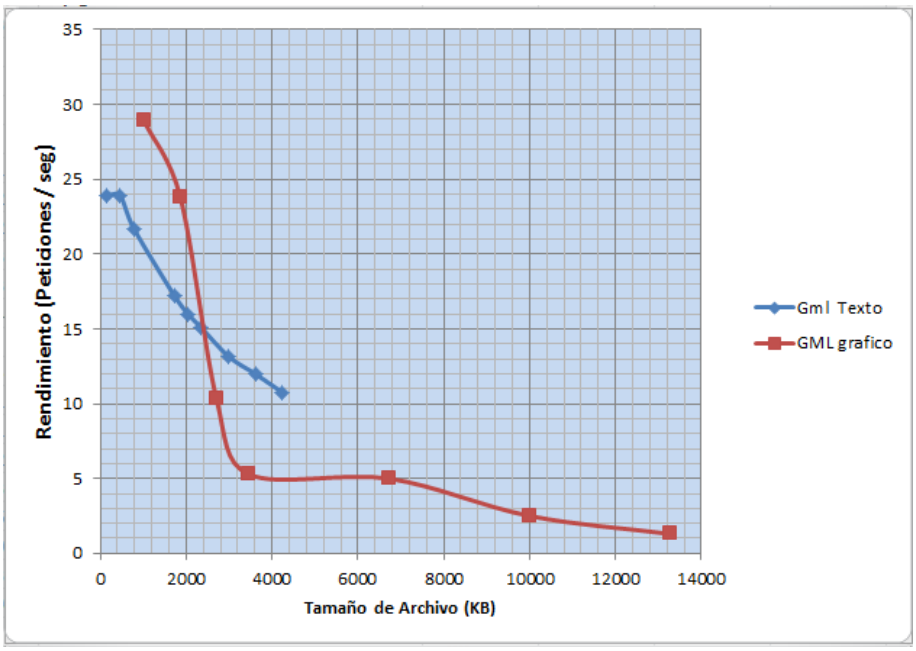


Figura 12 Cruce entre archivos donde el uno se manda solo el gráfico y el otro solo el texto, realizado con datos del DMQ

Segunda Hipótesis

Enunciado: Mientras mayor cantidad de información se tiene en los datos, la visualización del mapa es más rápida en OpenLayer que en MapServer y GeoServer.

CONCLUSIÓN:

OpenLayer con un servicio WMS, muestra un rendimiento muy satisfactorio e inclusive para archivos originales grandes el rendimiento sigue siendo satisfactorio, como se mencionó archivos ShapeFile sobre los 26 MB.

Tercera hipótesis

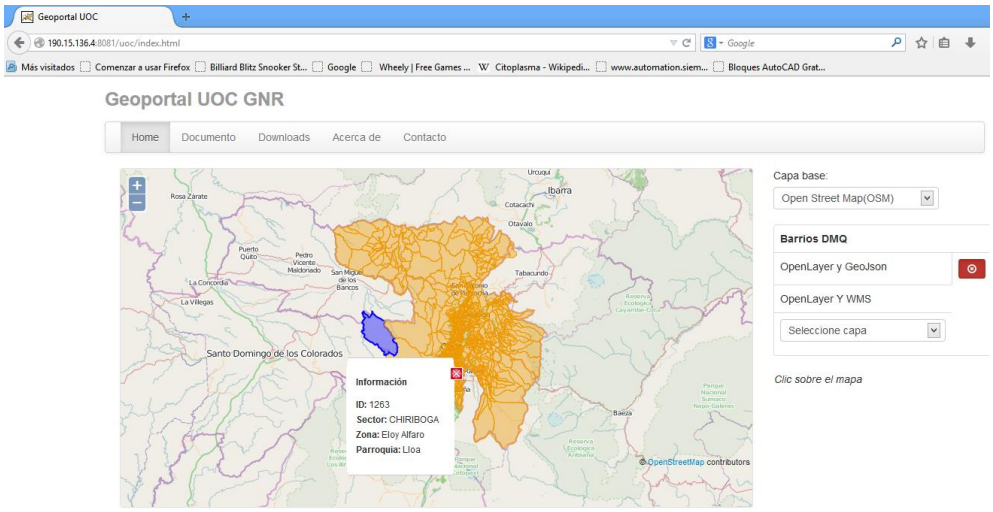
Enunciado: Para un mismo servidor de mapas, y un origen de datos dado, existe una relación logarítmica con los otros tipos de datos al momento de usarlo dentro del visualizador. Es importante destacar que estos orígenes de datos son:

- Shapefile,
- PostGis,
- GML
- GeoJson
- WMS

CONCLUSIÓN:

Como se mencionó en la Figura 8, fue posible establecer una relación exponencial entre los datos entregados al cliente de tipo GML y GeoJson, pero en un rango de valores y a partir de ahí esta relación fue lineal, por tanto si los datos son del mismo tipo, (por ejemplo vectorial), si es posible establecer una relación entre sus rendimientos, para un mismo origen de datos.

El rendimiento empleando OpenLayer con un servicio WMS resultó ser lo más satisfactorio. El geoportal diseñado muestra esto, así como se permite visualizar otro con una respuesta más lenta, considerando los origines de datos iguales. La dirección del geoportal es http://190.15.136.4:8081/uoc/index.html



7. REFERENCIAS

Anselin, L., Kim, Y., & Syabri, I. (2004). Web-based analytical tools for the exploration of spatial data. *Journal of Geograph System 6*, 197-218.

Carrasco Marimón, J., Belentani Cid, S., & Sanz Sa, J. (2013). Solución de alto rendimiento para el geoportal de turismo de la Comunidad Valenciana. *VI Jornadas de Software Libre.* Valencia.

Centro Virtual Cervantes. (2014). *Metodología cuantitativa.* Retrieved from http://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/diccionario/metodologiacuantitativa.htm

Cha, S.-J., Hwang, Y.-Y., Chang, Y.-S., Kim, K.-O., & Lee, K.-C. (2007). Integrating Ajax into GIS Web Services for Performance Enhancement. *International Conference on Multimedia and Ubiquitous Engineering (MUE'07)*, (pp. 668-673).

de Souza, B., Nunes, C., de Sousa, A., da Silva, E., Leite, F., & de Paiva, A. (Agosto 2005). On Performance Evaluation of Web GIS Applications. *Database and Expert Systems Applications, 2005. Proceedings. Sixteenth International Workshop on*, (pp. 497,501, 26-26). Copenhagen.

Flower, C. (2012, enero 21). *MapServer and GeoServer (and tilecache) comparison serving Ordnance Survey raster maps.* Retrieved from http://www.esdm.co.uk/mapserver-and-geoserver-and-tilecache-comparison-serving-ordnance-survey-raster-maps

Fu, P., & Su, J. (2010). *Web GIS: Principles and Applications.*

Garmin. (2010). *Manual de Usuario GPS Garmin Modelo GPSMAP 76 CSx.* Retrieved Febrero 2013, from GPS Manual: http://www8.garmin.com/manuals/GPSMAP76CSx_OwnersManual.pdf

Goodchild, M. (2012). The future of Digital Earth. *Annals of GIS 18:2*, 93-98.

Guan, W. W., Bol, P., Lewis, B., Bertrand, M., Berman, M. L., & Blossom, J. (2012). WorldMap – a geospatial framework for collaborative research. *Annals of GIS, 18:2*, 121-134.

Hong, J.-H., & Liao, H.-P. (2011). Incorporating visualized data completeness information in an open and interoperable GIS map interface. *Journal of the Chinese Institute of Engineers, 34:6*, 733-745.

Kresse, W. (., & Danko, D. (. (2010). *Springer Handbook of Geographic Information.* Heidelberg: Springer.

Laxton, J., Serrano, J.-J., & Arenas, A. (2010). Geological applications using geospatial standards – an example from OneGeology-Europe and GeoSciML. *International Journal of Digital Earth, 3:S1*, 31-49.

Lee, D., & Liang, S. (2011). Geopot: a Cloud-based geolocation data service for mobile applications. *International Journal of Geographical Information Science,25:8*, 1283-1301.

Li, S., Dragicevic, S., & Veenendaal, B. (2011). *Advanced in Web-based GIS, Mapping Services and Application.* London: CRC Press .

Maguire, D., & Longley, P. (2005). The emergence of geoportals and their role in spatial data infrastructures. *Maguire, D.J., and P.A. Longley. 2005. The emergence of gComputers, Environment and Urban Systems 29*, 3-14.

MapServer. (2014). *MapServer*. Retrieved from http://www.mapserver.org/es/index.html

Masó, J., Pons, X., & Zabala, A. (2012). Tuning the second-generation SDI: theoretical aspects and real use cases. *International Journal of Geographical Information Science, 26:6*, 983-1014.

Ministerio de Agricultura, Ganaderia, Acuacultura y Pesca. (2010). *SIG TIERRAS*. Retrieved septiembre 10, 2013, from http://servicios.sigtierras.gob.ec/

Moncrieff, S., West, G., Cosford, J., Mullan, N., & Jardine, A. (2013). An open source, server-side framework for analytical web mapping and its application to health. *International Journal of Digital Earth*.

Murray, S. (2009). *Estadística.* México: McGraw Hill.

Naranjo Martinez, A. (2014). Evaluación del rendimiento de los servicios WMS de Mapserver y Geoserver para la implementación IDE. *Revista Electrónica de Postgrados*.

Navas, G., & Prieto, P. (2011). Geoportales en el Ecuador. *La Granja*(2), 58-64.

Oates, B. (2008). *Researching Information Systems and Computing.* London: SAGE Publications Ltd.

OpenLayers. (2013). *OpenLayers.* Retrieved abril 29, 2013, from Wikipedia: http://es.wikipedia.org/wiki/OpenLayers

Piccinini, C., Smith, M., Hooke, J., & Hesketh, K. (2013). Bibliographic webmap: the Physical Landscape of Britain and Northern Ireland. *Journal of Maps*.

PostGis. (2013, enero 2). *PostGis 2.0 Manual.* Retrieved 5 30, 2013, from PostGis: http://postgis.net/docs/

Raymond, E. (1999). *The cathedral and the bazaar.*

Santo, M., & Lecumberry, G. (2005). *El proceso de medición: Análisis y comunicación de datos experimentales.* Retrieved from http://www.unrc.edu.ar/unrc/digtal/El_proceso_de_med.pdf

Scholten, M., Klamma, R., & Kiehle, C. (2006). Evaluating Performance in Spatial Data Infrastructures for Geoprocessing. *International Journal of Geographical Information Science*.

SEI. (2012). *Bangalore Urban Metabolism Project (BUMP) Geoportal.* Retrieved from http://www.urbanmetabolism.in/bump/

Shengru, T., Flanagin, M., Wu, Y., Abdelguerfi, M., Normand, E., & Venkata, M. (April 2004). Design strategies to improve performance of GIS Web services. *Coding and Computing, 2004. Proceedings. ITCC 2004. International Conference on , vol.2*, (pp. 444,448 Vol.2, 5-7).

Sistema Nacional de Información. (2014). *Acerca de SNI*. Retrieved abril 23, 2014, from http://sni.gob.ec/acerca-del-sni

Sletto, B., Muñoz, S., Strange, S., & Donoso, R. (2010). El Rincon de los olvidados: Participatory GIS. *Experiential Learning and Critical Pedagogy in Santo Domingo*, 111-135.

Valladolid, D. d.-U. (2011, junio 15). *Evaluacion y Explotacion de Sistemas Informaticos*. Retrieved from http://www.infor.uva.es/~miguelv/eesi/

Wikipedia. (2013, agosto 03). *GeoServer.* Retrieved from http://es.wikipedia.org/wiki/GeoServer

Wikipedia. (2013, 10 17). *JMeter.* Retrieved 2014, from http://es.wikipedia.org/wiki/JMeter

Wikipedia. (2013). *OpenLayers.* Retrieved abril 29, 2013, from Wikipedia: http://es.wikipedia.org/wiki/OpenLayers

Wikipedia. (2014, 4 1). *Firebug.* Retrieved 2014, from http://es.wikipedia.org/wiki/Firebug

Yang, B., Purves, R., & Weibel, R. (2007). Efficient transmission of vector data over the Internet. *International Journal of Geographical Information Science*.

Yang, C. (2005). Performance-improving techniques in web-based GIS. *InternationalJournal of Geographical Information Science, 19:3*, 319-342.

Yang, C., Wong, D., Yang, RX, Yang, R., Kafatos, M., & Li, Q. (2005). Performance-improving techniques in web-based GIS. *InternationalJournal of Geographical Information Science, 19:3*, 319-342.

Yang, J., Yang, C., Yu, Z., & Zhang, L. (Sept. 2004). Research on high-performance Web GIS system for map symbol dynamic editing and network publishing. *Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. vol.5* (pp. pp.2971,2974 vol.5, 20-24). IEEE International.

Yeşilmurat, S., & İşler, V. (1-07-2012). Retrospective adaptive prefetching for interactive Web GIS applications. *Geoinformática; Volumen 16*, 435-466.

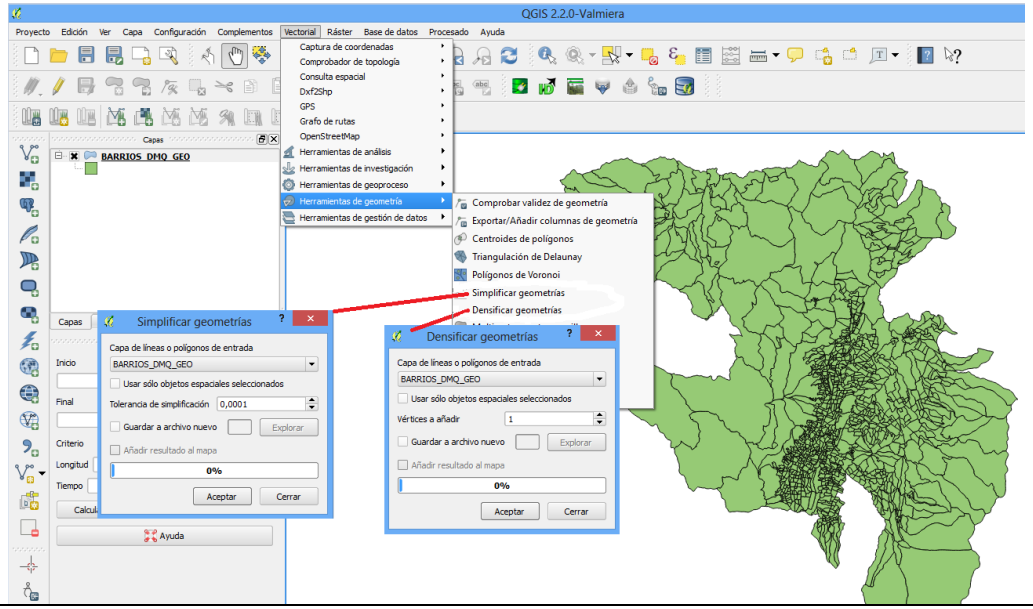
Zhang, T., Qiao, Q., & Wu, Q. (2009). Grid-Enabled High Performance GIS Web Portals. *Fifth International Joint Conference on Computational Sciences and Optimization, pp. 187-191, 2009 International Joint Conference on Computational Sciences and Optimization.*

ANEXO 1

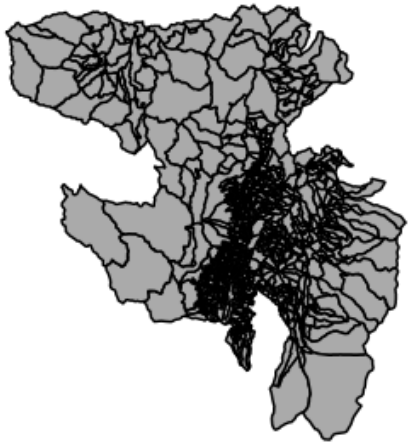
SOCIO ECONOMICO CON 76 COLUMNAS DE DATOS					
NOMBRE	DBF	SHP		No. Filas	No. Columnas
Pobreza (24 porvincias delEcuador)	1142,0	4354,0		936,0	76,0
pobrezamenosg (Menos Guayas)	1065,0	3899,0		874,0	76,0
pobrezamenosg2 (Menos El Oro)	994,0	3683,0		816,0	76,0
pobrezamenosg3 (Menos Manabí)	913,0	3348,0		749,0	76,0
pobrezamenosg4 (Menos Esmeraldas)	838,0	2998,0		687,0	76,0
pobrezamenosg5 (Menos Los Ríos)	807,0	2842,0		662,0	76,0
pobrezamenosg6 (Menos Sucumbíos)	768,0	2504,0		630,0	76,0
pobrezamenosg7 (Menos Zamora Chinchipe)	732,0	2403,0		600,0	76,0
pobrezamenosg8 (Menos Orellana)	706,0	2128,0		579,0	76,0
pobrezamenosg9 (Menos Morona Santiago)	639,0	1953,0		524,0	76,0
pobrezamenosg10 (Menos Pastaza)	618,0	1698,0		506,0	76,0
pobrezamenosg11 (Menos Loja)	514,0	1444,0		421,0	76,0

SOCIO ECONOMICO CON UNA SOLA COLUMNA DE DATOS					
NOMBRE	DBF	SHP		No. Filas	No. Columnas
gpobreza (24 provincias del Ecuador)	24,0	4354,0		936,0	1,0
gpobrezamenosg (Menos Guayas)	23,0	3899,0		874,0	1,0
gpobrezamenosg2 (Menos El Oro)	21,0	3683,0		816,0	1,0
gpobrezamenosg3 (Menos Manabí)	20,0	3348,0		749,0	1,0
gpobrezamenosg4 (Menos Esmeraldas)	18,0	2998,0		687,0	1,0
gpobrezamenosg5 (Menos Los Ríos)	17,0	2842,0		662,0	1,0
gpobrezamenosg6 (Menos Sucumbíos)	17,0	2504,0		630,0	1,0
gpobrezamenosg7 (Menos Zamora Chinchipe)	16,0	2403,0		600,0	1,0
gpobrezamenosg8 (Menos Orellana)	15,0	2128,0		579,0	1,0
gpobrezamenosg9 (Menos Morona Santiago)	14,0	1953,0		524,0	1,0
gpobrezamenosg10 (Menos Pastaza)	13,0	1698,0		506,0	1,0
gpobrezamenosg11 (Menos Loja)	11,0	1444,0		421,0	1,0

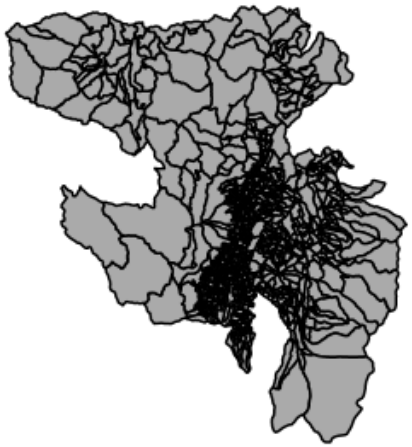
Archivo	s.01	s.001	s.0005	s.0001	DMQ Original	DMQ1	DMQ2	DMQ3
Formato Shape	Polígono	Polígono	Polígono	Polígono	Polígono	Polígono	Polígono	Polígono
Atributos	4	4	4	4	4	4	4	4
Número Filas	1269	1269	1269	1269	1269	1269	1269	1269
Archivo shp	173	280	406	935	3368	6645	9923	13201
Archivo dbf	81	81	81	81	81	81	81	81
Total shp y dbf	254	361	487	1016	3449	6726	10004	13282
Porcentaje shp	68,11%	77,56%	83,37%	92,03%	97,65%	98,80%	99,19%	99,39%
Porcentaje dbf	31,89%	22,44%	16,63%	7,97%	2,35%	1,20%	0,81%	0,61%



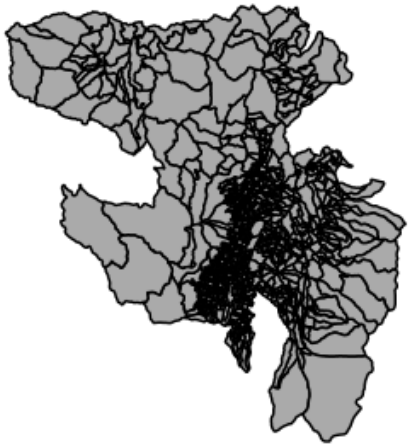
s.01 Tolerancia de simplificación 0.01



s.0005 Tolerancia de simplificación 0.0005

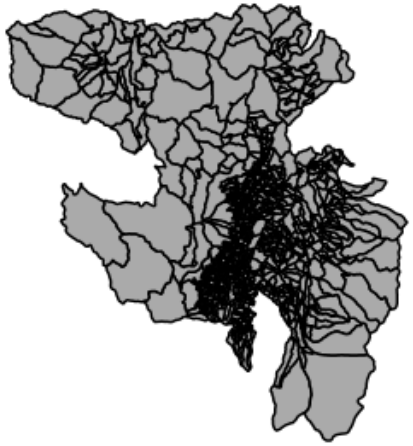


s.0001 Tolerancia de simplificación 0.0001



DMQ Archivo Original 211049 vértices

DMQ1 Vértices añadidos 1



DMQ2 Vértices añadidos 2