

## Parámetros, elementos y *triple store* a comparar.

**Objetivo:** Definir cuáles *triple store* y qué elementos serán considerados en la comparación.

**Resultados esperados:** Tabla comparativa de los *triple stores* similares a Apache Marmotta.

Los parámetros por considerar en el *benchmarking* a realizar después de la implementación del módulo de consultas federadas en el software Apache Marmotta son

- Tiempo de selección del *SPARQL endpoint*
- Tiempo de ejecución.
- Completitud de respuesta.

Los tiempos de selección y ejecución se medirán en segundos mientras que la completitud de respuesta en número de triplas que se recuperarán de las consultas federadas.

La definición y justificación de los parámetros propuestos se establecieron en un trabajo previamente desarrollado “Parámetros del *benchmarking*” donde se analizan 2 artículos *FedBench: A Benchmark Suite for Federated Semantic Data Query Processing* [1] y *Federated SPARQL Query Engines: Are Existing Testbeds Enough?* [2], en los cuales se propone cómo llevar a cabo un *benchmarking* en sistemas de consultas federadas. Tal como lo sugieren en [1], las consultas y los *datasets* serán adaptados con base a los *triple store* disponibles en:

- *Linked Data Cloud*<sup>1</sup>
- LODAtlas<sup>2</sup>

Las consultas por considerar son las que están disponibles en [1] en específico, las consultas de *Linked Data*. Cabe decir que estas consultas serán alteradas las consultas para darles un contexto espacial a los datos.

La infraestructura por usar durante el *benchmarking* es la siguiente:

- Computadora
  - o Sistema operativo: Unix-Linux Ubuntu.
  - o Apache Marmotta.
  - o Procesador: AMD A10-7850K ~3.7GHz.
  - o RAM: 8GB.
  - o Navegador Web: Google Chrome
- Conexión a Internet: 10Mbps.

---

<sup>1</sup> <https://lod-cloud.net>

<sup>2</sup> <http://lodatlas.lri.fr/>

El propósito del *benchmarking* es evaluar el desempeño del sistema completo para compararlo con otros similares. El sistema estará completo para nuestros propósitos después de implementar el módulo de consultas federadas.

Los *software* similares fueron extraídos de la tabla presentada en el trabajo “Comparar los diversos *triple store*” del actual alumno. Los *triple store* similares a Apache Marmotta, por alguna de sus funcionalidades, son lo que se observan en la tabla 1.

Tabla 1 Triple store similares a Apache Marmotta

Nombre	Consultas Federadas	GeoSPARQL	Status	Libre o pago	Large triple Store - Tamaño
Parliament	Si	Si - 2.7.4	Activo	Libre	No
AllegroGraph	Si	Si - Limitado a puntos en un cierto radio	Activo	Pago	Si - +1 Trillón
GraphDB	Si	Si	Activo	Ambos	Si - 17 billones Si - +39.8 billones
OpenLink Virtuoso	Si	Si	Activo	Ambos	Si - 1.08 Trillones
Oracle Spatial	si	No - tiene su equivalente en el release 2	Activo	Pago	Si - 50 Billones
Stardog	Si	Si	Activo	Pago	No
Marmotta	No	Si	Activo	Libre	No

*AllegroGraph*, *Oracle Spatial* y *Stardog* son sistemas que requieren un pago para ser usados mientras que *GraphDB*, *OpenLink Virtuoso* son sistemas que ofrecen un periodo de prueba y después adquirir una licencia para seguir usándolos.

Finalmente, el único *triple store* similar a Apache Marmotta es *parliament* al estar basado en el protocolo GeoSPARQL, ser libre y tener su propia versión de consultas federadas.

Por ende, los *triple store* a considerar por el hecho de ser libres en el *benchmarking* de consultas federadas geoespaciales son

- *GraphDB*.
- *OpenLink Virtuoso*.
- *Parliament*.
- *Apache Marmotta*.



## Referencias

- [1] M. Shmidt, O. Görlitz, P. Haase, G. Ladwig, A. Shwarte y T. Tran, «FedBench: A Benchmark Suite for Federated Semantic Data Query Processing,» de *International Semantic Web Conference*, Berlín, 2011.
- [2] G. Montoya, M. E. Vidal, O. Corcho, E. Ruckhaus y C. Buil-Aranda, «Benchmarking Federated SPARQL Query Engines: Are Existing Testbeds Enough?,» de *International Semantic Web Conferernce*, Berlín, 2012.