Proceedings of a pre-conference workshop of the 27th International Cartographic Conference:
Spatial data infrastructures, standards, open source and open data for geospatial (SDI-Open 2015)
20-21 August 2015, Brazilian Institute of Geography and Statistics (IBGE), Rio de Janeiro, Brazil

Using OS Software and Open Standards for Implementing a European SDI via Service Cascade

Lassi Lehto
lassi.lehto@nls.fi

Finnish Geospatial Research Institute, Geodeetinrinne 2, FI-02430 Masala, Finland

Keywords: service cascade, content integration, open source, spatial data infrastructure, web feature service

1. Introduction

The research described in this paper has been conducted in the context of a major EU project, called European Location Framework (ELF), initiated by EuroGeographics (EG), the co-operation organization of the European National Mapping and Cadastral Agencies (NMCAs) (EuroGeographics 2015). The ELF project aims at developing European-wide INSPIRE-compliant services from the geodata resources maintained by the EG's member organizations (European Location Framework 2015).

The ELF project started in March 2013 and will run for three years. The project has 30 participant organizations and 13 of them represent EU/EFTA member states as official NMCAs. Thus, the project has quite extensive spatial coverage extending from Finland to Spain and from Great Britain to Poland (see Figure 1).

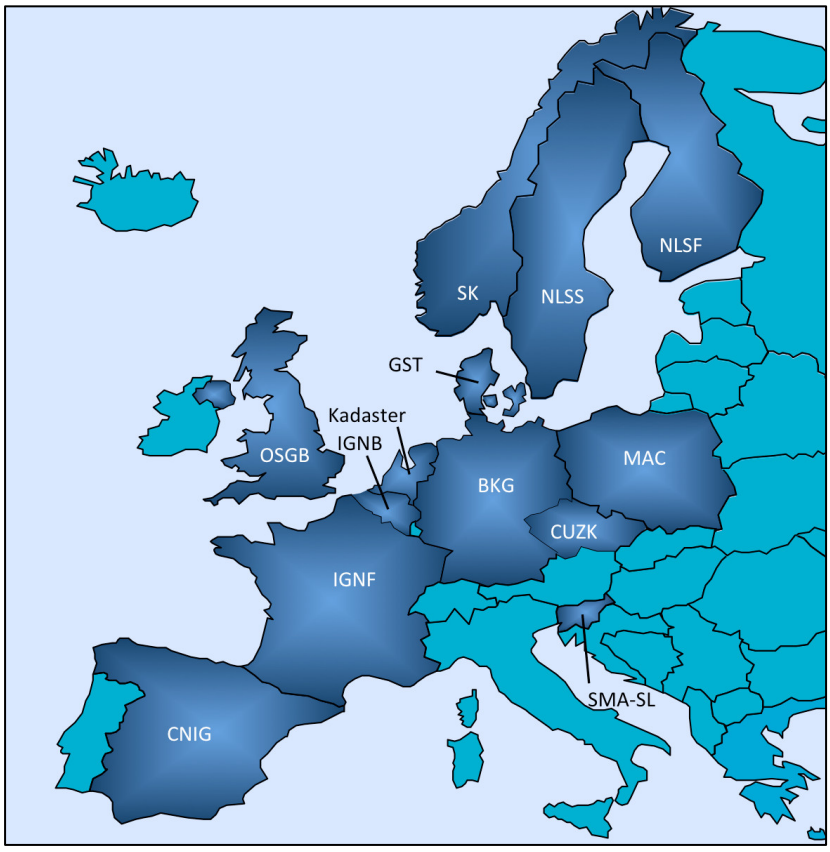

Figure 1. The countries participating in the European Location Framework project.

2. Aim of the research

The ELF project includes a work package specifically dedicated for data provision and service development. In this work package there is a subtask responsible for investigating the issues related to service cascade. The approach presented in this paper focuses specifically on the provision of European level Download Services (Web Feature Service, WFS) based on data services maintained by local NMCAs on country level (Vretanos 2011). The goal is to support the end user in accessing data content directly from national services. Thus, the cascading approach aims at carrying out real-time aggregation of content from a set of distributed national data sources (Lehto et al 2014).

One of the new challenges encountered when accessing national services from European-level applications is the need to introduce spatial integration capabilities to the service cascade approach. At the moment only thematic integration is supported in the existing cascade mechanisms of the Open Geospatial Consortium's Web Map Service standard. In this setup every single map layer is to be served by one and only one back-end service. When implementing cascaded integration over a set of national services, one has to resolve the problem of spatial query distribution and cross-border fusion of content. These are the biggest challenges in the further service cascade developments of the ELF project, too.

The goal of the cascaded services work in the ELF project has been to facilitate access to national download services providing master level datasets. The work has focused on direct access download services that conform to OGC's Web Feature Service (WFS) service specification. The aggregation layer developed by the cascaded services subtask of the project on top of the national services can be seen as a Cascading WFS implementation. Service cascade is not officially specified as part of the WFS standard. For this reason ELF Cascading WFS must be taken as an experimental test service.

3. Implementation

A content request coming from a client application is first analyzed by the Cascading WFS to determine, which national level services must be included into the process. Then the request is forwarded to the involved national services. The resulting datasets are merged together and finally returned back to the calling application. The analysis on the service inclusion is based on the bounding box of the query and on the requested feature types. The bounding box is overlaid on top of a dataset of national borders to determine, which countries the query overlap. The actual service inside the country is then selected depending on the requested feature type. ELF Cascading WFS does not support the full-fledged WFS service interface, but has certain significant restrictions, like the fact that the query must always contain a bounding box. The service has been developed as a Java Servlet-based web application. All the supporting software modules, like the DBMS for the country boundaries, the applied XML-processing libraries and the used software development tools are available as OS products. Most of the country level services are implemented using OS software implementations of the WFS standard.

The Cascading WFS component is running on the Amazon Cloud (EC2 small instance). The main tasks of the component is to decide to which national services the incoming WFS GetFeature –query must be forwarded, depending on the location of the query bounding box and the requested feature types. The bounding box of the query is overlaid with a national boundaries dataset stored in a separate PostgreSQL/PostGIS database. As a result, the

involved countries are listed. Then the requested feature types are checked to determine, which country level services have to be queried.

If only one backend service is involved in the Cascading query, the process is straightforward. The single service is accessed and the resulting dataset is returned to the calling application without further processing. If two or more services are involved, the Cascading WFS dispatches several parallel query threads to access the national services. The returning datasets are processed in the order they become ready. The datasets are merged into a single response message using a SAX-based XML-processing model.

The Cascading WFS has to create the root element (FeatureCollection) of the resulting dataset to include all required XML namespace declarations. In addition, the Cascading WFS fills in the boundedBy –element using the query's bounding box as the limits of the spatial extent. Then the contents of the individual background responses are written out as they become ready, taking away the FeatureCollection and boundedBy –elements. Finally the FeatureCollection element is closed by the Cascading WFS.

4. Limitations

The ELF Cascading WFS access interface has certain significant restrictions. Most importantly, it requires that a query must always have a bounding box defined. This can be presented as BBOX –parameter or as <fes:BBOX> -element inside either the FILTER –parameter of a GET-query or inside the <Filter> -element of a POST query. If there are several <Query> -elements in a single POST GetFeature request, they have to contain the same <fes:BBOX> (only the first one is taken into account). The ELF default Coordinate Reference System (CRS) is Web Mercator (EPSG:3857). This CRS can also be requested using the EPSG code 900913. Whether a given CRS is supported by the backend national level service varies from service to service.

Currently the GetCapabilities response message of the ELF Cascading WFS is configured manually. This could be based on dynamic GetCapabilities -querying of the backend national services. In any case there are certain limitations. All the backend services should support the same set of feature types, otherwise certain queries would result in an exception being returned. GetCapabilities response defines the spatial extent of each feature type by one single rectangle. As the spatial extents of the backend services are often distinct, the area definition in the GetFeature response becomes misleading. Because the other parameters, like the list of supported CRSs, must be composed as the greatest common denominator, some possibilities of the services remain unusable.

Feature identifiers must be globally unique for the Cascading WFS to work. In the case of the INSPIRE feature identifiers this is handled properly by using well-defined namespace mechanisms. Some of the ELF national services already apply an URI-based namespace label, which automatically ascertains global uniqueness. However, this becomes crucial in the case of the XML ID-typed gml:id –attributes used in feature and geometry elements. These values must also be prepended with the namespace label to keep them unique in the content aggregation phase, or some other mechanisms have to be employed.

5. Conclusion

The real-time content aggregation that is tested in the ELF project is based on the principle of Cascading WFS. ELF Cascading WFS does not support the full-fledged WFS service

interface, but has certain significant restrictions, like the fact that the query must always contain a bounding box.

Altogether 32 national download services have been included into the ELF Cascading WFS (Poland 5, Sweden 4, Norway 3, Czech Republic 10, Finland 4, France 1, Spain 5). These services provide access to 79 different feature types.

Acknowledgements
The ELF project has been funded by the European Commission, Grant Agreement No 325140.

References

EuroGeographics 2015, Home Page. http://www.eurogeographics.org Accessed 9 June 2015

European Location Framework 2015, Project Home Page. http://www.elfproject.eu Accessed 9 June 2015

Lehto, L., Latvala, P. and Kähkönen J. 2014, Service Cascade as a Means for Pan-European Access to National Geodata Content CASE: European Location Framework. The Sixth International Conference on Advanced Geographic Information Systems, Applications and Services, ”GEOProcessing 2014”, Mar 23 – 27, 2014, Barcelona, Spain, CD-ROM.

Vretanos P. A. [ed.] 2011, OpenGIS Web Feature Service 2.0 Interface Standard. http://portal.opengeospatial.org /files/?artifact_id=39967 Accessed 9 June 2015