

# Cascading Geospatial Services for Integration of Authoritative National Datasets – CASE: European Location Framework

Lassi Lehto, Pekka Latvala, Jaakko Kähkönen

Finnish Geodetic Institute, PO Box 15, FI-02431 Masala, Finland

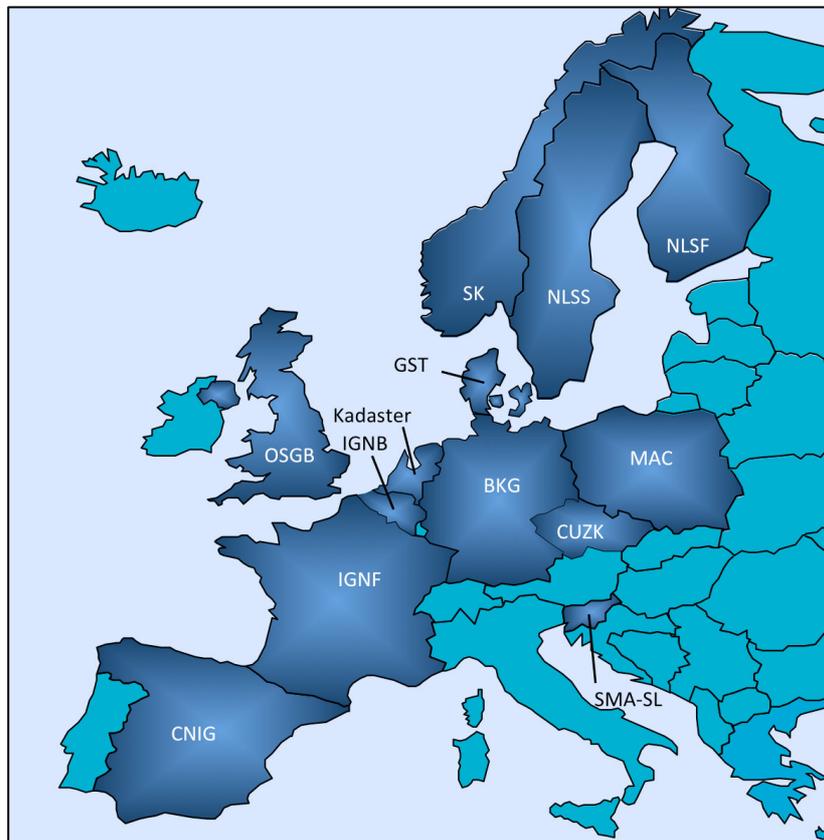
**Abstract.** The concept of service cascade can be seen as a solution for the data aggregation needs set forth by the basic INSPIRE principle, according to which European level Spatial Data Infrastructures (ESDIs) can be built on top of national level SDIs. The research described in this paper has been conducted in the context of a major EU project, called European Location Framework (ELF). The main goal of the work has been to facilitate European level access to national master datasets using direct access download services that conform to Web Feature Service (WFS) specification. Content aggregation is based on the Cascading WFS approach. Service cascade is not formally specified in the WFS standard, thus the ELF Cascading WFS is presented as an experimental implementation with some significant limitations. Altogether 32 national download services have been successfully included into the ELF Cascading WFS, providing access to 79 different feature types from 8 INSPIRE themes.

**Keywords:** Service Cascade, Web Feature Service, European SDI

## 1. Introduction

Providing European-wide access to geospatial data resources held on national level is the ambitious goal of the Infrastructure for Spatial Information in the European Community (INSPIRE) process and other similar integration initiatives (European Commission 2007). Service cascade is presented in this paper as a solution for facilitating access to national content to support Pan-European applications. The paper presents the process of providing a centralized access point to geospatial data, requested from several national INSPIRE-compliant Download Services (European Commission 2009).

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 based on geodata resources maintained by the EG's membership. The ELF project started in March 2013 and will run for three years (European Location Framework 2015). The project has 30 participant organizations, 13 of them representing EU/EFTA member states as official NMCAs. Thus, the data resources accessible by the project have quite extensive spatial coverage across Europe, see *Figure 1*.



**Figure 1.** The countries participating in the European Location Framework project. The label indicates the National Mapping and Cadastral Agency (NMCA) representing each country.

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 (Lehto et al. 2014). The approach presented in this paper covers this development and focuses specifically on the provision of European level download services based on data services delivering content on national level. The approach is based on the principle of a Cascading Web Feature Service (WFS) (Vretanos 2011).

The *Section 2* introduces the concept of service cascade and its implementation in the download services of the ELF platform. The *Section 3* provides a more detailed description of the ELF Cascading WFS module and *Section 4* discusses its most significant limitations. The paper ends with concluding remarks in *Section 5*.

## **2. Service Cascade**

### **2.1. Principle**

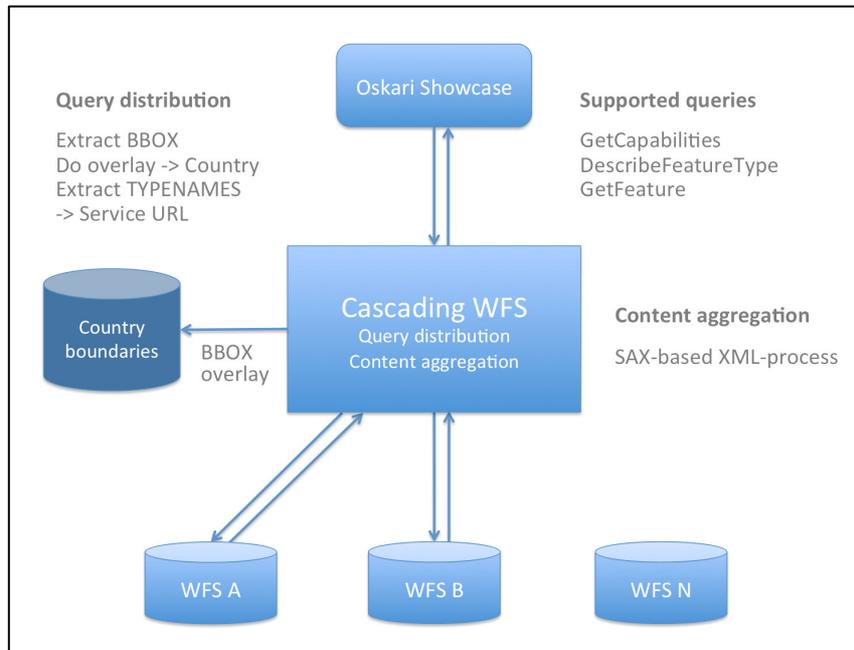
Concept of service cascade is evaluated in this paper as the solution for data aggregation requirements. According to the service architecture model of the Open Geospatial Consortium (OGC), the basic idea in service cascade is that a service end point can be configured as a content source for another service, actually making this latter service a client of the source service (Open Geospatial Consortium 2015). Service cascade can be regarded as an implementation of the basic INSPIRE principle, according to which European level Spatial Data Infrastructures (ESDIs) can be built on top of national level SDIs. This can be seen as the most cost-effective way for building services on multiple levels of local administration as well. In the past, service cascade has been studied for instance in the context of metadata services integration (Deng and Wu 2010).

### **2.2. Query Distribution**

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 geospatial content.

The main tasks of the ELF Cascading WFS 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. The ELF Cascading WFS approach is depicted in the *Figure 2*.



**Figure 2.** Cascading WFS service architecture of the ELF platform.

### 2.3. Content Aggregation

The goal in the ELF platform development is to support the end user in accessing geospatial data content both as aggregated on European level, and directly from national services. Thus, the cascading approach aims at supporting real-time aggregation of content from a set of distributed national data sources. One of the new challenges encountered when accessing national services from European-level applications is the need to introduce spatial integration capabilities to the traditional service cascade approach. At the moment only thematic integration is supported in the existing cascade mechanisms of the OGC service implementations (Web Map Service, WMS) (Beaujardiere 2006). In this setup every single map layer is served by one and only one back-end service.

If only one backend service is involved in the Cascading WFS query, the process is straightforward. The single service is accessed and the resulting dataset is returned to the calling application without further processing. However, 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 (Simple API for XML 2004); see *Figure 2*.

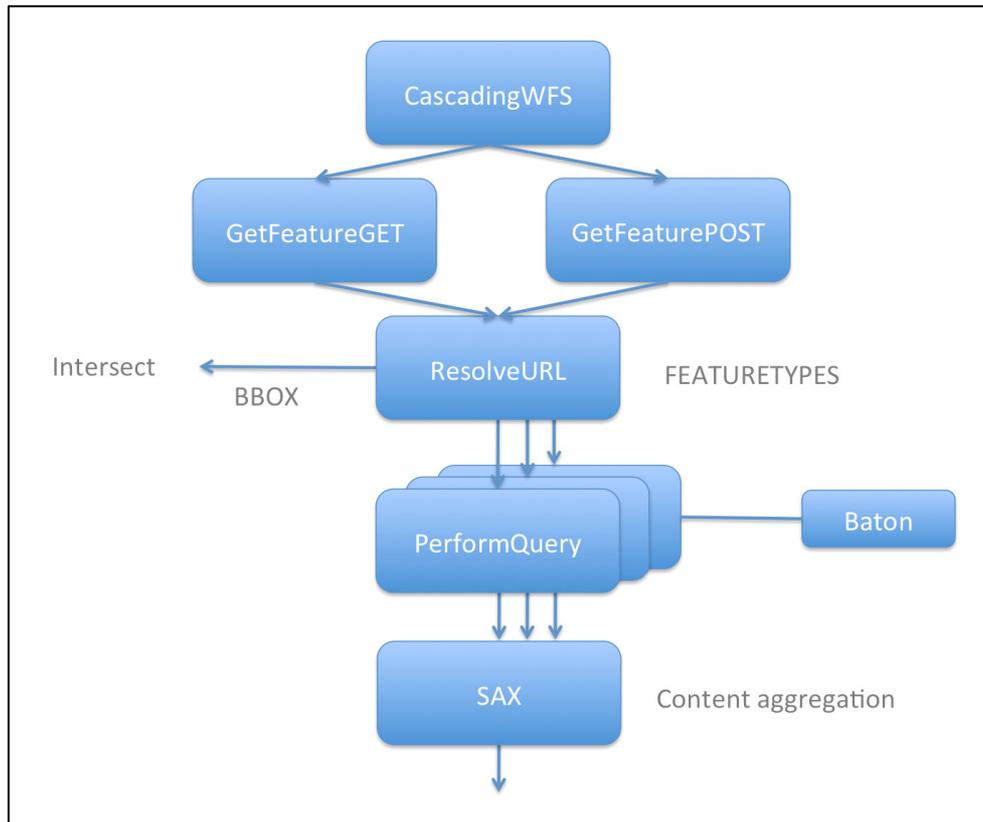
The Cascading WFS has to create the root element (FeatureCollection) of the resulting dataset to include all required XML namespace declarations. In addition, the Cascading WFS fills in the boundedBy –element using the bounding box of the query as the limits of the spatial extent. Then the contents of 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.

## **2.4. Client applications**

The ELF Cascading WFS functionality has been tested with a wide set of test queries using a standard Web browser. The POST-queries have been tested from browser using a service-side module that forwards the queries to the service interface. In addition, the Cascading WFS has been successfully tested with the QGIS client application (QGIS 2015). Integration with a showcase application based on the Finnish Open Layers-derived map library, Oskari, has been partially tested (Oskari Platform 2015); see *Figure 2*.

## **3. Implementation**

The Cascading WFS implementation is running on the Amazon Cloud (EC2 small instance) as a Java Servlet on top of a Tomcat Web application server platform. Backend national services are queried by simultaneously dispatching a set of threaded query processes. The responses are returned to the client application in the order they become ready. A baton-based approach is used to prevent the parallel threads from writing to the response stream at the same time. The internal modules of the ELF Cascading WFS are depicted in *Figure 3*.



**Figure 3.** Internal structure of the ELF Cascading WFS implementation.

#### 4. Limitations

The ELF Cascading WFS access interface has certain significant restrictions. Most importantly, it requires that the query must always have a bounding box defined. This can be presented as a BBOX –parameter or as a <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 Coordinate Reference Systems (CRSs) that are supported by the Cascading WFS include Web Mercator (EPSG:3857 or EPSG:900913), WGS84 (EPSG:4326), ETRS89 (EPSG:4258) and Lambert Equal Area (EPSG:3035). Country boundaries dataset is stored in the database in these CRSs and thus the overlap with the query bounding box can be tested only using these four CRSs. The ELF default 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. The 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 GetCapabilities response becomes misleading. Because the other parameters, like the list of supported CRSs, must be defined 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 a well-defined namespace mechanism. Some of the ELF national services already apply a 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 kept unique in the content aggregation phase, for instance by prepending them with the namespace label, 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.

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.

Altogether 32 national download services have been included into the ELF Cascading WFS. These services provide access to 79 different feature types from 8 INSPIRE themes.

## References

- Beaujardiere J de la [ed.] (2006) OpenGIS Web Map Server Implementation Specification. [http://portal.opengeospatial.org/files/?artifact\\_id=14416](http://portal.opengeospatial.org/files/?artifact_id=14416) Accessed 9 April 2015
- EuroGeographics (2015) Home Page. <http://www.eurogeographics.org> Accessed 9 April 2015
- European Commission (2007) INSPIRE Directive. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:EN:PDF> Accessed 27 March 2015
- European Commission (2009) COMMISSION REGULATION (EC) No 976/2009 of 19 October 2009 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards the Network Services. <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02009R0976-20101228&from=EN>. Accessed 27 March 2015
- European Location Framework (2015) Project Home Page. <http://www.elfproject.eu> Accessed 27 March 2015
- Lehto, L., Latvala, P. and J. Kähkönen (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.
- Open Geospatial Consortium (2015) Home page. <http://www.opengeospatial.org> Accessed 27 March 2015
- Oskari Platform (2015) Home Page. <http://oskari.org> Accessed 9 April 2015
- P. A. Vretanos [ed.] (2011) OpenGIS Web Feature Service 2.0 Interface Standard. [http://portal.opengeospatial.org/files/?artifact\\_id=39967](http://portal.opengeospatial.org/files/?artifact_id=39967) Accessed 27 March 2015
- QGIS (2015) Home Page. <http://www.qgis.org/en/site/> Accessed 9 April 2015
- Simple API for XML (2004) Home Page. <http://www.saxproject.org> Accessed 9 April 2015
- Y. Deng and Q. Wu (2010) Research on the harvest and cascade of catalogue service in GeoGlobe Service Platform. Proc. 18<sup>th</sup> International Conference on Geoinformatics: GIScience in Change (Geoinformatics 2010) Peking University, Beijing, China, June, 18-20, 2010.