Service Cascade as a Means for Pan-European Access to National Geodata Content

CASE: European Location Framework

Lassi Lehto, Pekka Latvala and Jaakko Kähkönen
Department of Geoinformatics and Cartography
Finnish Geodetic Institute
Masala, Finland
lassi.lehto@fgi.fi, pekka.latvala@fgi.fi, jaakko.kahkonen@fgi.fi

***Abstract*—Providing European-wide access to geospatial data resources held on national level is the ambitious goal of the INSPIRE process and other similar integration initiatives. Service cascade is presented in this paper as a solution for facilitating access to national content to support Pan-European applications. The paper presents as an example the process of creating a map tile cache of geospatial data requested from INSPIRE-compliant Download Services. Cascading service architecture is found to be a useful approach in the cache seeding process. The work has been conducted in the context of a major EU-project.**

Keywords-service cascade; content integration; map tile cache; cross-border applications

I. Introduction

According to the fundamental INSPIRE (Infrastructure for Spatial Information in the European Community) principles, Pan-European geodata-related services are supposed to be built on top of national Member State services [1]. The same basic principle holds in the national context, where state level offerings are in many cases dependent on the availability of content and services on local government level. The traditional solution for this issue is to collect content from the local sources into a centrally managed data store and provide higher level services from that aggregated resource. In the rapidly changing world, this approach has certain drawbacks. One of the most serious problems is that the aggregated data sets fast become outdated. Various methods have been developed for incremental update of service databases from the original sources, but satisfactory results are yet to be achieved. Recently, Open Geospatial Consortium (OGC) has initiated a work related to the topic with the title GeoSynchronization, but software implementations of this work are still largely missing [2].

The concept of service cascade is evaluated in this paper as the solution for the data aggregation problem. The basic idea in service cascade is that a service can be configured as a content source for another service, actually making this latter service a client of the first service. Service cascade can be regarded as the most fitting implementation of the basic principle, according to which higher level Spatial Data Infrastructures (SDIs) can be built on top and making use of lower level SDIs. This can also be seen as the most cost-effective way for building services on multiple administrative hierarchy levels. Previously service cascade has been studied for instance in the context of metadata services integration [3].

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) [4]. The ELF project aims at developing European-wide INSPIRE-compliant services from the geodata resources maintained by the EG's membership.

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.

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 covers the first phase of this development and focuses specifically on the provision of European level view services based on data services delivering pre-aggregated content.

In Section II the origins of the service cascade approach in standardization are described. In Section III the concept of map tile cache is discussed, together with the related standardization work. Section IV describes the phase one cascading service architecture adopted in the ELF project. Section V details the contents of the tile cache resulted from the service cascade work in the project and Section VI explains the challenges faced and solutions found in this process. Section VII concludes the paper, together with a short discussion about plans on the further development of the service cascade approach in the ELF project.

II. Service Cascade

The idea of service cascade has been present in the OGC's service specifications since the early days. Already the Web Map Service (WMS) version 1.0.0 introduced the functionality to define another WMS as the content source for a certain layer of information that a given WMS serves

 ISBN: 978-1-61208-326-1

[5]. In addition to just combining the content offerings of other services with its local map layers, the cascading WMS can provide other functionalities like coordinate system transformations, new image formats, or even more advanced image manipulation or analysis operations. Service cascading can also be used to make use of the WMS standard versions that are not supported by the original back-end service.

In the Styled Layer Descriptor (SLD) specification, the service cascading approach has been further elaborated by introducing a concept of Component WMS, a service instance that can be freely associated with any back-end content source [6]. Typically, these kinds of service combinations represent the case, in which a Web Feature Service (WFS) source is accessed by a WMS instance to provide a visual representation of the geospatial data content available in the background service.

III. MAP TILE CACHE

A. General

The concept of map tile cache has been introduced as the means for improving the performance of map delivery to the Web-based client applications. The idea is to pre-render at least the most frequently requested parts of the back-end data store in a form of a regularly gridded map tile cache. The most typical map tile image size is 256 * 256 pixels.

When a client application sends a request for a certain area, the tile service first checks, whether those map tiles already exist in the cache or not. If they are already in the cache, the tiles are returned right away to the calling application. If the tiles have not yet been rendered, the content for these tiles are requested from the back-end service, rendered, stored to the cache and returned to the client. This way the cache will automatically get populated for the parts that are actually needed and requested by clients. On the other hand, map tile cache implementations typically provide functionalities for fully pre-filling the cache from the data store. This process is called the feeding of the cache.

B. Tile Cache standards

Various standards have been developed as the specification for the map tile cache access interface. These include WMS Tile Caching (WMS-C) of the Open Source Geospatial Foundation (OSGeo), and the most recent Tile Map Service Specification (TMS) of the OSGeo foundation [7]. The most official one is the Web Map Tile Service (WMTS) 1.0.0 standard of the OGC [8]. In addition, the map tile service might also support the traditional WMS interface, but this requires that the service is able to assemble a single image, corresponding to an arbitrary query window, from the stored tiles.

Another important standard related to a map tile cache is the grid used to delimit the individual tiles. This grid must be bound to a Coordinate Reference System (CRS) and declare the origin and the extent of the area covered. A complete tile cache specification also includes information of the used tile size in the ground scale. Because the cache typically supports various different display scales, the contents of the cache will be generated on several zoom levels, effectively making the cache a hierarchical, pyramid-like structure. The tile size must be defined for each of the zoom levels. Often this is expressed as list of resolutions, i.e., the grid cell sizes in ground units.

IV. SERVICE ARCHITECTURE

In the first phase, the cascaded services approach in the ELF project is realized as a map tile cache containing theme-wise visualizations of the pre-aggregated Pan-European data sets Euro Global Map (EGM), Euro Regional Map (ERM) and Euro Boundary Map (EBM). Thus, in this stage the cascading service approach is specifically adopted in the tile cache seeding process. The cascaded service architecture consists of several components (Figure 1). Data content is provided by a set of WFS service instances maintained by the German NMCA, BKG, in Leipzig. The data sets have been pre-transformed to the INSPIRE-compliant form and are thus available according to the themes and feature classes, as defined in the INSPIRE data specifications.

The tile cache is implemented using Open Source software products and deployed as an Amazon Web Services (AWS) -based cloud service. The service consists of the following main components: 1) a WMTS-compliant service implementation (based on a product called MapCache), 2) a WMS service instance (MapServer), responsible for the map rendering process, 3) a seeding script, customized in the project and utilizing the '`mapcache_seed`' process.

Because the used map rendering engine only supports WFS version 1.0.0 and simple features data model as the back-end source, a custom-built middleware module had to be introduced. Over time this module has been further developed to also support other functionalities. The middleware component runs on the service platform of the Finnish Geodetic Institute (FGI) in Finland.

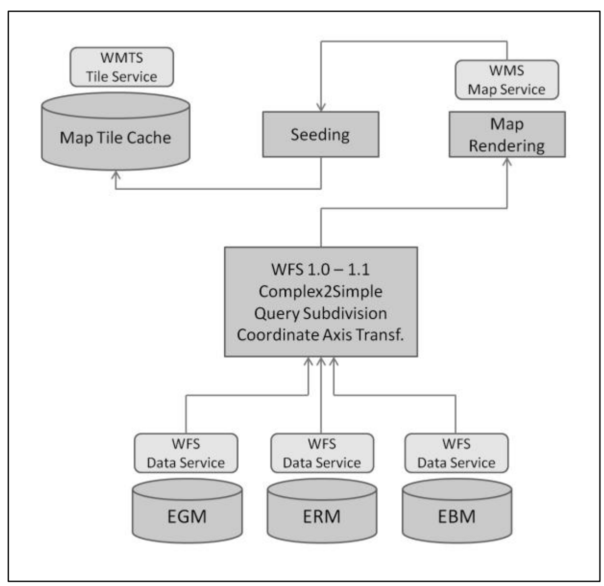

Figure 1. Service architecture of the ELF cascading services, phase 1.

Functionality of the middleware module covers four main tasks. Firstly, it takes care of the differences between the WFS version 1.0.0, supported by the map rendering engine as the source data source, and the WFS version 1.1.0 available in the back-end services. Secondly, it transforms the complex INSPIRE schemas into a simpler form that can be digested by the rendering engine. Thirdly, it performs a recursive query window subdivision, needed to resolve the timeout problem of the back-end services in case of very large polygonal features. Finally, the middleware component flips the coordinate axis order of the incoming data sets. This is required, as the treatment of CRS-related details is different in WFS 1.0.0, when compared with that in WFS 1.1.0.

V. CONTENTS OF THE TILE CACHE

At the moment the ELF map tile cache contains six INSPIRE-compatible map layers listed in Table I.

TABLE I. ELF MAP TILE CACHE CONTENTS

INSPIRE Map Layer
AU.AdministrativeBoundaries
HY.PhysicalWaters.Watercourse
HY.PhysicalWaters.StandingWater
HY.PhysicalWaters.LandWaterBoundary
TN.RoadTransportNetworks.RoadLink
TN.RailTransportNetwork.RailwayLink

All these map layers are cached in five zoom levels that correspond to the grid sizes standardized in the INSPIRE process (100 km and 10 km), enhanced with some additional levels to facilitate smooth zooming in map viewing applications. The five zoom levels, the corresponding grid sizes, resolutions and map display scales are shown in Table II.

TABLE II. ELF MAP TILE CACHE ZOOM LEVELS

Level	Grid size [km]	Resolution [m]	Appr. Scale
0	200 * 200	781.25	1 : 3 M
1	100 * 100	390.625	1: 1.4 M
2	40 * 40	156.25	1 : 500 000
3	20 * 20	78.125	1 : 280 000
4	10 * 10	39.0625	1 : 140 000

An exemplary display of the current map tile cache contents is shown in Figure 2. This includes all currently available map layers in the area of Southern Portugal. The map shown is based on Euro Global Map content in approximate scale of 1 : 3 M.

The available map layers are displayed using the INSPIRE-defined simple default styles. All layers are rendered with transparent background. This is to ease overlaying of individual map layers on top of various different types of geospatial content by third-party applications. Place names are missing from the current content offering of the ELF map tile cache, but can be integrated form the EuroGeoNames service, if needed [9].

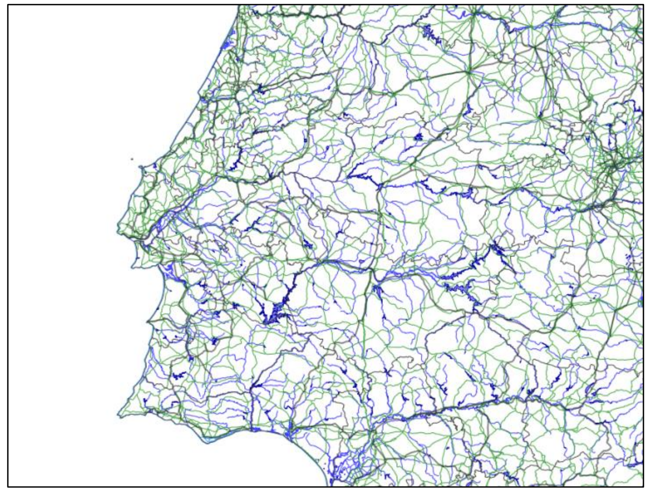

Figure 2. An exemplary map display containing all the available map layers of the ELF map tile cache, from Southern Portugal

The content in the ELF map tile cache is rendered in the Lambert Equal Area CRS (EPSG:3035) and covers the whole continental Europe, having the spatial extent of N_{min}:1300000, E_{min}:2600000, N_{max}:5500000, E_{max}:6000000.

VI. PERFORMANCE AND FEASIBILITY CONSIDERATIONS

User-experienced performance is one of the main motivations for the development of map tile cache-based service architectures. Map tile creation is often run as a pre-processing step, called seeding of the cache. This process can be partial, i.e., carried out only in areas having most demand for map visualizations. In this approach, the areas requested less frequently will be accessed from the original source and rendered on-the-fly. User-experienced performance is naturally poorer in those areas.

In the case of the ELF project, a complete pre-seeding of the tile cache is performed. Full seeding enables the use of very poor performance data sources while keeping the user-experienced responsiveness of the system well inside the limits established in the INSPIRE regulation on service performance. In the ELF project the approach is to make use of WFS-based data download sources for map rendering process. Even complicated feature geometries can be successfully processed in this approach.

However, in the case of the most voluminous geometries, additional facilities had to be developed. In the ELF project the main reason for this was the fact that the source data services would time out on queries that involve several very complicated polygonal features, like those present in the 'AdministrativeUnit' and 'StandingWater' feature classes. Even most careful adjustment of the corresponding timeout thresholds in the middleware service and the rendering engine would not resolve this issue. Special functionality was developed in the middleware service component to tackle the problem.

When encountering the back-end service time out situation, the middleware component divides the query window into four sub-windows and re-initiates the data query as four sub-queries. If the four requests are successful, the middleware component integrates the corresponding responses as a single result data set that is returned to the rendering process. If any of the sub-queries fails due to the time out of the back-end service, the subdivision function is called recursively until a successful result is achieved. Duplicate features created in this process are ignored as they do not harm the rendering process.

VII. CONCLUSION AND OUTLOOK

The map tile cache approach used in the ELF project enables good end user experience of the map application performance, even though low performance data source services are used. The approach is an implementation of the service cascade architecture, in which a service acts as a client to other services. The created map tile cache serves content according to the principles established in the INSPIRE process. Thus, content is divided into individual map layers and made available in the CRS and grid, as defined in the INSPIRE process. Map tiles have a transparent background to facilitate integration with other data sources.

In future, the cascading services subtask of the ELF project will focus on providing cascading functionality for the data download operations. The goal is to support the end user in accessing data content available both as pre-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. In this setup every single map layer is designated 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 map images. These are the biggest challenges in the further service cascade developments of the ELF project, too.

Cascading of data level services also raises new challenges, because some European NMCAs do not allow caching of data content outside their organizational firewalls. In addition, many NMCAs provide access to data download only in file-based forms.

One of the preliminary plans concerning the cascading of file-based data services is to provide access to content through an interactive map user interface. According to INSPIRE, file download services are supposed to be available as Atom feeds. From the ‘coverage’ metadata element in the Atom feed, one can retrieve the spatial extent of the data in the file and display it as a rectangle on the map. This would facilitate the downloading task considerably.

A more advanced approach would be to facilitate access to file-based download services through the individual tiles of the map tile service. This would require overlap of the tile extent with available file content extents and on-the-fly selection of the portion of each file that is inside the map tile. Functionalities for this have not yet been designed or developed in the ELF project.

Another possible functionality for a cascading service in the context of data download is to convert the INSPIRE-compatible Geography Markup Language (GML) -encoded content coming from the back-end download services into a more Web-friendly format, like JavaScript Object Notation (JSON). This reformatted content could then be made available via some modern Web application-oriented access interface.

ACKNOWLEDGMENT

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

REFERENCES

[1] European Commission, INSPIRE Directive, at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:EN:PDF [accessed: 2014-01-22]

[2] Open Geospatial Consortium, Specifications Working Group on GeoSynchronization Home Page, at: http://www.opengeospatial.org/projects/groups/geosyncswg [accessed: 2014-01-22]

[3] Y. Deng and Q. Wu, “Research on the harvest and cascade of catalogue service in GeoGlobe Service Platform,” Proc. 18th International Conference on Geoinformatics: GIScience in Change (Geoinformatics 2010) Peking University, Beijing, China, June, 18-20, 2010.

[4] European Location Framework (ELF), Project Home Page at: http://www.elfproject.eu [accessed 2014-01-22]

[5] Open Geospatial Consortium, “OpenGIS Web Map Server Interface Implementation Specification, Revision 1.0.0”, April 2010, at: http://portal.opengeospatial.org/files/?artifact_id=7196 [accessed: 2014-01-22]

[6] Open Geospatial Consortium, “Styled Layer Descriptor profile of the Web Map Service Implementation Specification, Version 1.1.0”, June 2007, at: http://portal.opengeospatial.org/files/?artifact_id=22364 [accessed: 2014-01-22]

[7] Open Source Geospatial Foundation, Home page at: http//www.osgeo.org [accessed: 2014-01-22]

[8] Open Geospatial Consortium, “OpenGIS Web Map Tile Service Implementation Standard, Version 1.0.0”, at: http://portal.opengeospatial.org/files/?artifact_id=35326 [accessed: 2014-01-22]

[9] L. Lehto, P. Latvala and J. Kähkönen, “An Implementation of the OGC’s WFS Gazetteer Service Application Profile, CASE: The EuroGeoNames Central Service Renewal,” Proc. The Fifth International Conference on Advanced Geographic Information Systems, Applications and Services, ”GEOProcessing 2013”, Feb 24 – Mar 1, 2013, Nice, France.