Web Feature Services,
Considerations for CGDI Government Partners

Version 1.0

This report was made possible by the Natural Resources Canada, Earth Sciences Sector, Geomatics for Connecting Canadians Program.

This report was prepared by the team from the AMEC Earth and Environmental Technology Business Unit, led by Chris DeJager. The team is based in Ottawa Canada and is one of 90 AMEC Earth and Environmental offices world-wide. AMEC Earth and Environmental has over 2000 employees designing, delivering and supporting environmental services to both the private and public sector. The Technology Business Unit is a centre of excellence for enterprise information management and spatial information systems within AMEC.

Table of Contents

1. Executive Summary

In reviewing WFS in the context of the Atlas of Canada's desire to further integrate into the data management and dissemination processes of CGDI partner programs, there are four main areas of consideration.

1. WFS digital rights and business process support;
2. WFS security and performance;
3. WFS as a service used for data dissemination between data stakeholders at the Municipal, Provincial, and Federal levels; and
4. WFS as a potential technical aggregator of feature and attribute data for service to client applications.

Each area of consideration for the use of WFS needs to be considered independently to ensure any limitation in one area does not overshadow the overall potential use associated CGDI endorsed specifications.

Many organizations have initiated implementation of WFS and have ended up serving GML files instead. As a result there are few true WFS implementations active for comparison or review. Those implementing WFS indicate performance issues, GML version changes and limited client application adoption of WFS capabilities as reasons why WFS has seen a slower adoption than WMS.

WFS services which have been largely successful, such as the Geographic Names of Canada WFS, are focused on point geometry types with small numbers of returned records and attributes. In addition WFS and WMS used within the same application based on the appropriate user functions and relative data scale may provide the content rich data of WFS within the performance expectations of users.

Overall WFS has capabilities that are best applied to data transfer and processing applications where access times to data are expected to be slower but the resulting analysis product is of great value. In considering the use of WFS within the broader CGDI it is best to relate its use back to specific functional processes where exchange of vector data and attributes results in maintenance of distributed data nodes.

WFS's are not a high priority to average end users. WFS/GML remains a capability utilized by GIS professionals and students with a technical knowledge of the use of GIS.

Recent developments in simplifying the GML schema, GMLsf, will allow vendors and users to operate on a static GML schema and in turn begin to overcome some of the performance limitations of WFS 1.1.0 GML 3. As this new GML standard takes hold CGDI government partners will be able to move forward and take advantage of WFS in three main areas:

1. Providing data for aggregation into National Frameworks, for example GeoBase, as part of an automated mechanism;
2. Managing data by acquiring it from source asynchronously to the request/response processes for current client applications;

3. Develop linked attribute databases with separately stored geographic features linked at runtime or as part of an update process;
4. Investigate end client uses of WFS and incorporate functionality back into CGDI client applications, therefore addressing users' needs.

WFS represents a standards based approach to dealing with vector data and attributes therefore opening up the data management capabilities of the Atlas of Canada to integrate and participate with partner organizations.

2. Introduction

Natural Resources Canada (NRCan) is committed to leveraging each self managed data source, such as the Atlas of Canada, National Topographic Database, and GeoBase, to serve their specific client needs. Each of the organizations responsible for these data are committed to utilizing the CGDI endorsed standards and specifications for spatial data interoperability and web services. By ensuring adherence to a common set of standards each organization can implement best practices for the maintenance of their data and applications while leveraging resources of other standards adopting organizations.

As part of this consistent implementation of standards, Natural Resources Canada has implemented CGDI standards-based technologies and is currently providing data through the Web Map Service (WMS) standard for the Atlas of Canada. Other organizations are now leveraging the Web Feature Service (WFS) standard in addition to WMS .Natural Resources Canada in accommodating this new capability is assessing the potential role for WFS and the different business and technical impacts.

2.1 Purpose

The purpose of this document is to review the WFS standard, how it has been addressed by other government organizations, review the business and technical implications in engaging WFS within the broader CGDI and to provide recommendations.

This document is not an assessment of current implementations of WFS by government organizations, and does not compare vendor product suitability for use in WFS.

2.2 Contributors

The following list of people contributed to the content of this document through participation in interviews.

Person	Organization	Contact
Doug O’Brien	Natural Resources Canada, GeoAccess Division	(613) 947-1287
Jean-Pierre Lemieux	Natural Resources Canada, Centre for Topographic Information - Sherbrooke, CTIS	(819) 564-5600 (232)
Pierre Desjardins	Natural Resources Canada, Centre for Topographic Information - Sherbrooke, CTIS	(819) 564-5600 (297)

Brad Ruth	Statistics Canada, Geography Division	(613) 951-6433
Catherine Caldwell	Statistics Canada Geography Division	(613) 951-6587
Peter Rushforth	Statistics Canada Geography Division	(613) 951-3890
Colin MacDonald	GeoNova, Nova Scotia and Municipal Relations	(902) 424-5281
Peter Schut	Agriculture and Agri-Food Canada NLWIS IM/IT Support	(613) 759-1874
Mike Adair	Natural Resources Canada, GeoSystems	(613) 947-1342

2.3 References

GeoDRM WG, Digital Rights Management (http://www.opengeospatial.org/groups/?iid=129)

Geography Mark-up Language v2 – (http://portal.opengeospatial.org/files/?artifact_id=1034)

Geography Mark-up Language v2.1.1 – (http://portal.opengeospatial.org/files/?artifact_id=1108)

Geography Mark-up Language v2.1.2 – (http://portal.opengeospatial.org/files/?artifact_id=11339)

Geography Mark-up Language v3 – (https://portal.opengeospatial.org/files/?artifact_id=7174)

Geography Mark-up Language v3.1.1 – (http://portal.opengeospatial.org/files/?artifact_id=4700)

Geography Mark-up Language Simple Feature Profile – (http://portal.opengeospatial.org/files/?artifact_id=11266)

OWS Common Implementation Specification - (http://portal.opengeospatial.org/files/?artifact_id=6324)

OWS-3 Data Aggregation Service (DAS), OGC Document 05-120 (http://www.opengeospatial.org/initiatives/?iid=162)

Web Feature Service 1.0.0 – (https://portal.opengeospatial.org/files/?artifact_id=7176)

Web Feature Service 1.1.0 – (https://portal.opengeospatial.org/files/?artifact_id=8339)

Web Feature Service (Transactional) – (https://portal.opengeospatial.org/files/?artifact_id=7176)

3. Web Feature Service

3.1 What is WFS?

A Web Feature Service (WFS) is an interface to a server that allows a client application to call for geographic features and attributes. When the server responds to the request for geographic features and attributes it sends a definition of the data in vector format, including attributes, allowing the client application to render the vectors and display on the map (Figure 1). The volume of vector and attribute data returned is a function of the size of the area selected as well as the density and complexity of the vectors.

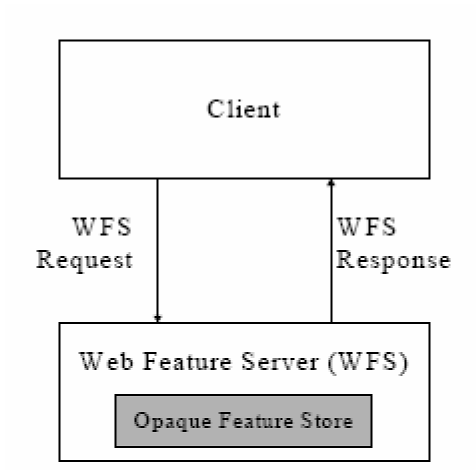

Figure 1: WFS Request/Response

3.2 How does it work?

The actions performed in responding to a WFS request are outlined in Figure 2. The client application makes the request to the identified WFS server requesting data. The WFS server selects the data and processes it into an open readable format called Geography Mark-up Language (GML) which is based on the extensible mark-up language (XML). XML is as international standard for file information exchange and GML is a recognized ISO and OGC standard. GML allows for the description of geography and associated attributes in an open manner such that any application that can read XML can read GML.

The GML is generated based on an associated schema referenced by the WFS. The GML is streamed back to the client application where it is validated against the identified schema and rendered into the map.

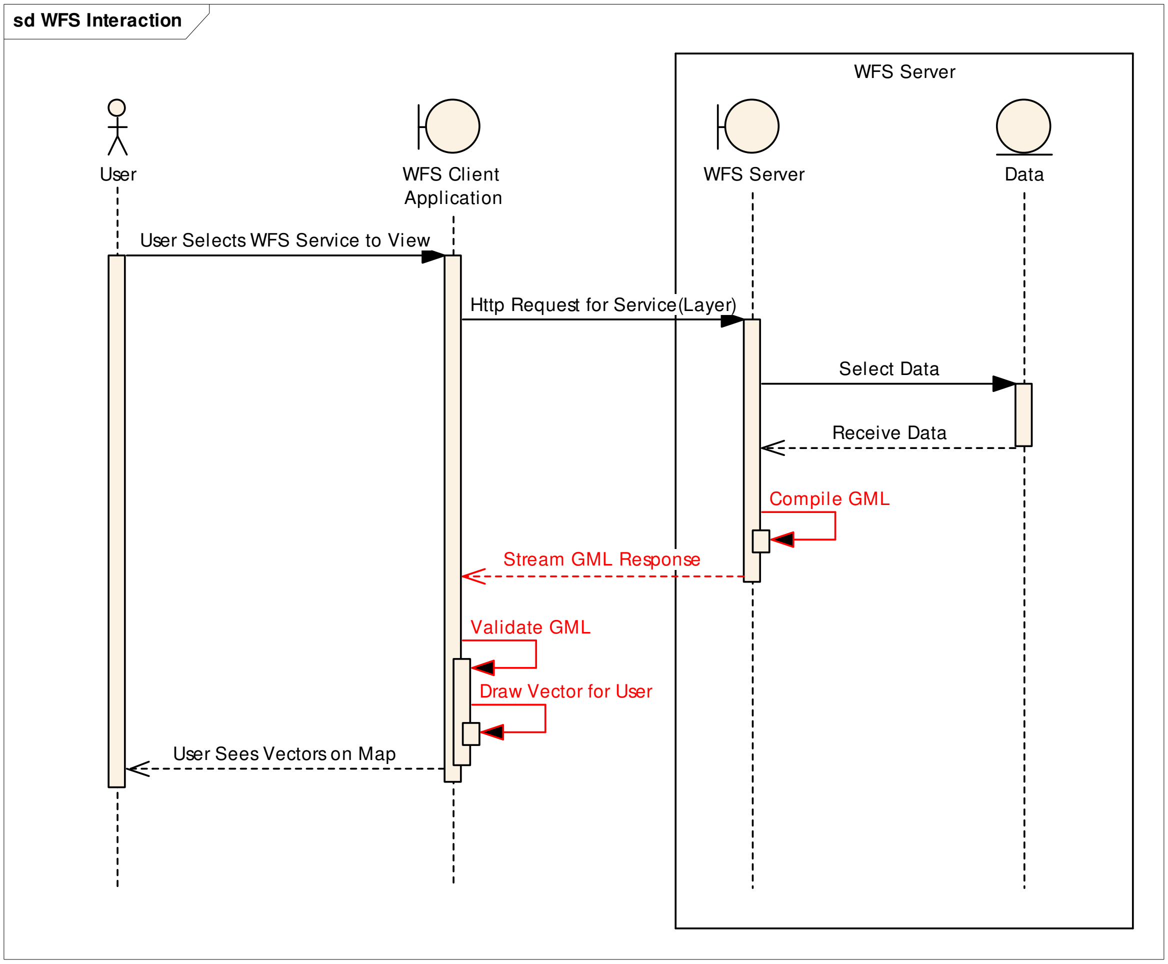

Figure 2: WFS Interaction

3.3 Differences with WMS

WMS has been a foundation service of CGDI. Effort to utilize WMS pervasively and to performance tune the service creates a benchmark against which all other user facing CGDI services are measured. The marked difference in performance between WMS and WFS require a detailed understanding of the sequence of activities performed by WMS to compare with the activities of WFS. By Understanding the sequence of activities of WMS attention can be focused on the additional activities WFS performs in response to a request.

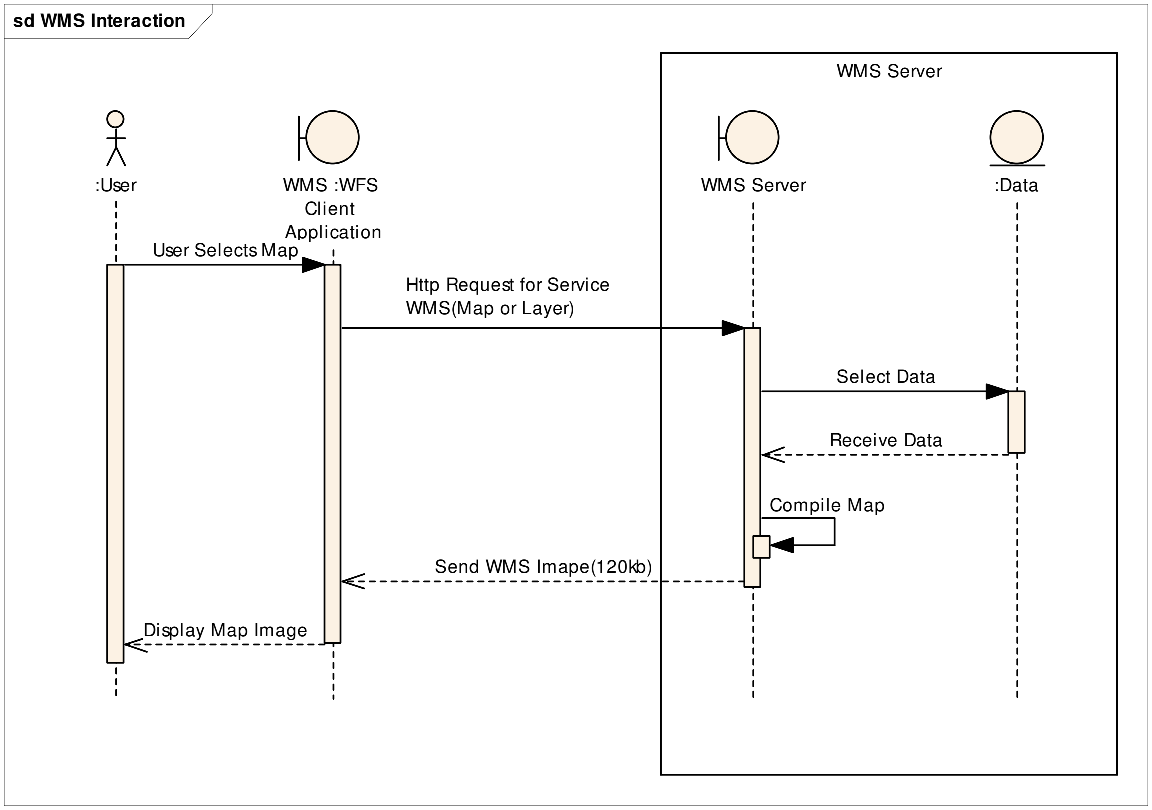

Figure 3: WMS Interaction

WMS interaction consists of an http request, a data pull, a compiling of the map, and the sending of the map image as the response (Figure 3). The resulting map image is always a set size and does not change with the amount of data portrayed on the map. As a result the WMS service has a performance consistency that can be linearly improved through additional hardware resources or software configurations.

3.4 WFS Process Performance

By understanding how a WFS call is broken down, processes that differ from WMS highlight steps that slow the service cycle time when compared to WMS. Specifically, all of the sequence steps highlighted in red in Figure 2 are steps that are performance bound by the amount of data to being returned. The volume of vectors being returned impacts each process by increasing the processing time resulting in a logarithmic increase in overall service time and resources. Twice as many vectors returned is not double the service time, it is more.

Calculation of exact performance expectations has many variables including;

- The number of vectors and their complexity add to the size of the GML file and increases the rendering time of the vectors on the client application;

- The number of attributes associated with each vector increases the size of the GML file;
- Attributes being referenced using XLink impact the processing time when rendering the GML file as the application may not move through the GML file linearly; and
- The increasing complexity of GML schemas impacts time required to perform the validation processes by the application.

All of these variables need to be accommodated by client applications to be a “universal” WFS viewer. The industry to-date has not been able to reconcile all of these factors into an efficient client viewer application and therefore different client viewers have different user experiences with the same WFS.

Comments on performance of WFS do not take into consideration impacts of the WFS server hardware and the network connectivity bandwidth between the client application and server. Performance is the key weighting by users of WFS when considering use.

3.5 Versions, Operations and Service Classes

As of March 2006, there are two versions of Web Feature Service (WFS), 1.0.0 and 1.1.0, which are endorsed standards within the CGDI. WFS was first adopted as an implementation specification, an endorsed specification by the Open Geospatial Consortium Inc. (OGC), May 7, 2002 as version 1.0. The second version of WFS, version 1.1 was endorsed May 3, 2005.

There are some fundamental differences between WFS 1.0.0 and WFS 1.1.0. At a high level WFS 1.1 has added a new operation GetGMLObject and a corresponding class of service, XLink. The full operations and service class comparisons are in Table 1 and Table 2.

Operations	WFS 1.0	WFS 1.1.0
GetCapabilites	✓	✓
DescribeFeatureType	✓	✓
GetFeature	✓	✓
GetGmlObject		✓
Transaction	✓	✓
LockFeature	✓	✓

Table 1: WFS Operations Comparison

Operations are the functions that can be initiated through the WFS call. These operations support the internal functions of transaction and query processing.

GetCapabilities returns a description of the feature types supported by the WFS and what which operations are supported for each feature type. This basically describes the complete “capability” provided to uses of the specific service.

DescribeFeatureType returns a description of the feature type that has been requested.

GetFeature allows the client application to specify a feature(s) to return and the attributes to be returned with the feature(s).

GetGmlObject provides a capability to place references within feature definitions references, internally of externally to the GML data elements. This allows for the separation of feature geography and attributes until they are merged at the time of rendering.

Transaction allows modification functions to be performed on features. The Transaction operation includes the ability create, update, and delete features and/or associated attributes.

LockFeature creates a lock on identified features. This may be used during a transaction operation to inhibit modification from another request or to remove the possibility of modifications to select features.

Class	WFS 1.0	WFS 1.1.0
Basic	✓	✓
Transaction	✓	✓
XLink		✓

Table 2: WFS Classes Comparison

Based on the WFS operations for each version there are service classes that can be defined.

Basic WFS implements the GetCapabilities, DescribeFeatureType and GetFeature operations. The resulting service would be read-only WFS.

XLink WFS supports all operations of a basic web feature service and in addition would implement the GetGmlObject operation for local and/or remote XLinks, and offer the option for GetGmlObject operation to be performed during GetFeature operations. The resultant response GML would contain a link to the resource. Also the resource value could be retrieved as the response GML is generated resulting in a fully contained dataset within the GML within the client application.

Transaction WFS supports all the operations of a basic web feature service and in addition it would implement the Transaction operation. Transaction WFS could optionally implement the GetGmlObject and LockFeature operations.

In addition to the operations additions between WFS 1.0.0 and WFS 1.1.0 there is also a change in the base GML schema used. The following schema versions of GMLare required as a minimum for the identified WFS version.

	WFS 1.0	WFS 1.1.0
Shall use GML schema version	2.1.1	3.1.1

Table 3: WFS to GML schema version

3.6 Geography Mark-up Language (GML)

GML is the heart of WFS. WFS and GML have evolved independently since GML can be read natively as a file format by GIS applications and not solely served through WFS. As a result GML versions identified in WFS specifications are not the most recent versions of GML available.

Specification	Date
GML 1 (R)	2000-05-12
GML 2	2001-02-20
GML 2.1.1	2002-01-14
GML 2.1.2	2002-08-19
GML 3	2003-01-29
GML 3.1	2004-04-19
GMLsf	2006-03-17

Table 4: GML Version History

Note: (R) identifies that the version did not become an implementation specification, it was replaced by subsequent versions before approval.

GML 2 deals with defining geographic features using the simple feature description. This means that GML 2 features are 2D and are simple points, lines and polygons. GML 3 marks a shift to defining features as 3D objects, complex non-linear geometry, temporal properties, and 2D topology. As a result the GML 3 base schema is 8 times as large as the GML 2. A schema is a description of all feature types and property elements to be found in the GML feature file. Schema processing is a required function in validation and rendering of GML.

GML application schema's for GML 3.1.x are backwards compatible to versions GML 2.1.2 and GML 3.0 0. This means that with little or no change the older schemas can be used to produce a GML 3.1.x schema. With WFS 1.0.0 supporting a required GML version of 2.1.1 (Table 3) then the schemas currently servicing WFS 1.0.0 capabilities are invalid and would need to be re-generated to support a WFS 1.1.0. This also means that to support both versions of WFS two separate application schemas will need to created and maintained in support of all potential user client application service requests.

According to the OGC registered products web page there are only a few products that implement WFS 1.1.0. There are over 350 products registered and very few even use GML 3, the majority are GML 2 compliant. In an effort to move the vendor community and therefore the user community to WFS 1.1.0 there has been a new GML specification generated that allows for simple feature definitions but maintains many of the advances of GML 3.

Simple Feature GML (GMLsf) has a greatly reduced schema from GML 3. It can be considered a schema subset of GML 3 and implements 90% of what service providers needed. This is meant to be a static schema so that anything generated using GMLsf will be interoperable. The complexity of GML 3 allowed vendors to add elements that were not necessarily compatible with all client software. GMLsf will provide this standard implementation and therefore eliminate the key barrier to implementation of WFS 1.1.0.

3.7 Client Application and Server Negotiation

As part of the WFS specifications there is a process of “negotiation” between the client and the server. This negotiation is meant to allow future versions of the implemented WFS standard to respond to multiple clients requesting different versions of the WFS interface.

Negotiation occurs when a client application, as part of its http request, requests a version of WFS the server does not support. The server’s response is to offer the next lowest possible version to the client application. The client application either accepts the offered version, making a subsequent request using this version, or it requests a different version. The server continues to accept or provide an alternate version. This “negotiation” continues in a sequence that moves from the most recent to the oldest version available on the server. If the client application and the server cannot agree on a version then the service request is ultimately refused and the client application does not receive the data.

Currently this is only an issue if a WFS is only 1.1.0 and does not respond to 1.0.0 requests. Since most WFS implementations, over 90%, are the base version 1.0.0 the negotiation is as simple as a yes or no. Most client applications do not “negotiate” they simply request a version (1.0.0) and if they do not get it then the “negotiation” is over. With the implementation of WFS 1.1.0 using GMLsf eagerly anticipated by service providers and the fact that the schema’s supporting WFS 1.0.0 and WFS 1.1.0 are not backwardly compatible the client applications will need to accommodate great variability in WFS services. This incompatibility between WFS 1.0.0 and WFS 1.1.0 has most likely been the key obstacle in implementing WFS client applications that deal with all WFSs.

4. Case Studies of WFS Implementation

In order to capture the experience of implementing WFS and GML selected organizations were interviewed. It is important to note that all interviews were conducted before March 12th, 2006. This date is significant because GMLsf was adopted as a standard on that date and some case study comments reflect the anticipation of this standard.

4.1 Statistics Canada

Statistics Canada (STC) is a federal government leader in providing data in GML format. The September 2005 release of the Road Network File (RNF) signified a crucial step for STC in contributing to a larger interoperable distributed CGDI architecture. WFS and GML are reasonable service capabilities for STC since the data being provided contains attributes desired for geocoding or analysis, not just symbolization.

STC set the goal of releasing standards compliant road network data and chose the CGDI endorsed GML as the format that would provide the broadest potential use within the GIS community. GML allows STC to meet CGDI standards compliance and be client application agnostic. The resulting RNF GML files are compliant to the GML 2.1.2 standard.

An assessment of the GML standard and resulting files sizes was conducted to ensure that performance using specific GML client applications under normal computing conditions, provided sufficient accessibility and usability. STC does not provide a WFS interface to their GML data. Instead users access the STC website and download the appropriate GML files clipped to provincial boundaries.

Why has STC not implemented a WFS to provide GML data? Simply. STC has calculated that the overhead associated with the service would have a greater impact on the web infrastructure of STC compared to the downloading of compressed pre-packaged GML files. In providing the complete file for an area through a download mechanism removes the need for users to query the server to generate and regenerate the same file content. In addition, with an annual data release, users do not generally require a live service.

From a business perspective STC has dealt with digital rights and access through the use of an end user license agreement (EULA) presented to the user online. The user is required to agree to this EULA before they download any data.

STC looks to a future where data can be exchanged between provincial and municipal road network maintainers and federal value add custodians. The transactional capability of WFS would allow municipalities through the provinces, to submit changes to the RNF editor team for integration and validation.

4.2 GeoNova

GeoNova is the province of Nova Scotia’s portal for spatial information. Over the past year several projects have allowed GeoNova to provide 4 public WFS and 2 internal WFS services. In addition, independent of WFS, files are available in GML format for 5 products and 30-40 individual layers

The GML structure generated through the WFS and the file creation process are slightly different although both “meet” GML 2 standard. The inherent addition of elements into the GML schema, supporting additional functionality for specific vendor client applications, may interfere with the overall reading of the GML by client applications for which the elements are not supported.

GeoNova is implementing WFS in an innovative manner from the standpoint of performing transformations of the data before rendering in GML. These transformations are on the attributes of a feature or the feature geometry, and are consistently applied as part of the extract, transform and load process to generate downloadable GML files. The result is that different features and attributes are being linked to produce static GML files for download by users.

GeoNova is supporting acquisition of data from municipalities for aggregation within the GeoNova data framework. This currently is completed by the municipalities supplying data files, as shapefiles, directly to GeoNova for re-work into the framework. Looking forward, GeoNova envisions a transactional model in place that would ingest delta changes from municipalities to the province as a WFS service. With the complexity of GML 3.0 GeoNova considers this to be an ambitious goal. That being said they are encouraged by the new simple feature GML schema as the schema to facilitate the transaction processing required to meet the vision.

From a business context GeoNova is evolving their infrastructure to implement the SAP portal enterprise and through this technology will provide a level of service security and access management. To date GeoNova has supported open access to available web services (WMS and WFS) but with the desire to move to a more transactional model there will be a need to implement a new access/security model for “peer-to-peer” transactional WFS.

Currently access to data and services is gated through a free registration and access process. Users with accounts are actively presented with an end user license agreement (EULA) that needs to be accepted before access to data is provided. GeoNova has taken the steps to make the EULA pervasive throughout their distributed data partnerships. This means that upon accepting the one EULA the user is licensed for use of all GeoNova and partner data holdings and services. This makes accessing Nova Scotia data easy, managed and seamless.

4.3 National Topographic Database

The National Topographic Data Base (NTDB), maintained by Natural Resources Canada, is not currently using WFS or GML in its production environments. There have been a few enquiries by private companies and OGD’s requesting access to data through WFS but there is no clear indication from users that WFS is becoming a required data service.

The NTDB viewer application Toporama has integrated with the Atlas of Canada Web application . NRCan is following the lead of the GeoBase Program as it defines the direction to be taken with adoption and adherence to services standards expected of suppliers of GeoBase Portal layers.

NRCan echoes the concerns of GeoNova on performance of WFS and GML. The deciding factor for implementation of WFS/GML will reside with the requirement for this service for GeoBase Portal layers.

NRCan has identified the distributed data management model as a priority. Currently CTI-S utilizes validation scripts to process data received by jurisdictions for incorporation into their national coverages (NTDB, NRN). These validation processes are currently based on the shapefile format but GML or XML updates could provide greater standardization and a move to a more automated validation and processing. WFS would be a mechanism to feed data into the validation process.

There are business concerns with ensuring metadata and copyrights are maintained with data once it leaves GeoBase Portal, Atlas of Canada, or distribution of the NTDB. NRCan is waiting for a coordinated policy digital rights to be applied to the larger federal government (Atlas of Canada, GeoBase Portal, Statistics Canada, etc…).

Recent developments in NRCan have provided the opportunity to harmonized policy in this area through the new Data Management and Dissemination Branch.

4.4 Geobase Portal

The GeoBase Portal implemented by Natural Resources Canada (NRCan), as a contribution to the national GeoBase Program, is providing a common access point to national data frameworks within Canada. Currently there are seven layers available at a national coverage including but not limited to hydrography and roads. The GeoBase Portal currently supplies data in file format for download, as GML or shapefiles, and provides a WMS. Currently there is no WFS implementation, but there is interest from the national Geobase program participants in future provision of this service.

GeoBase maintains a simple and effective end user licence agreement for its data layers and provides minimal security to services. Acknowledgement of this end user license is required prior to accessing the GeoBase WMS. This approach ensures users are informed of copyrights and data service limitations while providing open access to services.

4.5 Comments

It is interesting to note that for the 21 WFS's registered with CGDI/CEONet not one is WFS version 1.1.0. The fact that organizations implementing WFS are not upgrading to version 1.1.0 might be in part be due to the complexity of GML 3 and also the client application/WFS server negotiation that may not be available within client applications. In this case it is best to use the base WFS version.

The GeoBase Portal is a natural focal point for service additions, such as WFS, related to national coverages of data. Since the GeoBase Portal relies on supplier aggregation of jurisdictional data implementation of WFS needs to be reviewed in relation to the supplier services and the use of WFS as a facilitator of automated aggregation. In addition the supplier aggregation process being driven by provincial/municipal, data may in the future necessitate a distributed access (de-centralized) portal architecture.

Although organizations have had success in implementing WFS the Geonames WFS is an indication that a well organized and focused service capability can result in a very good performing WFS for general public consumption.

5. WFS Implications for CGDI Federal Partners

CGDI Government Partners face tremendous challenges in considering the use of WFS. These challenges are not only technical in nature but have business implications as well. These challenges apply equally to Federal, Provincial and Municipal partners. The following section outlines some of the major challenges and their potential implications to CGDI Government Partners.

5.1 Business Implications

5.1.1 Service Level Commitments

With the adoption of interoperable standards for accessing geographic information government contributors to CGDI have been challenged to meet the ever increasing service level expectations of the user community. Users of web services provided through CGDI have come to expect instantaneous access to information. In addition the web services are expected to be highly available, even during peak usage.

WFS represents a challenge since the current infrastructure supplying typical web services at today's service levels cannot support the same level of service for WFS.. The resulting WFS service is not a 2-fold increase in service time but a minimum of 5 to 10- fold increase in service time depending on the WFS standard implemented and the feature types described within the WFS. WFS point services such as Geographic Names of Canada have acceptable performance capabilities.

With applications within one organization relying more and more on data and services provided by another organization, service levels are critical to the functioning of end user applications. Service levels and availability become the key factors to end users and application developers when choosing CGDI standards.

5.1.2 Service Security

There are two main architectures for securing web services. The most common architecture treats web services as an object and the network security controls the user's access to objects through the use of roles. Roles are groups of functionality applied to a user allowing roles to be modified to apply to large groups of users without individually managing each user's discreet functions profile. This common architecture is currently IT best practices for security.

A new architecture is emerging based on the security being set by the objects within the network. Services would have inherent security access interfaces and the individual services would authenticate users as part of the access protocol. This new service security architecture allows for decentralization of user authentication and services. This new architecture provides users' access to services that "exist" and are not within a registry. Service "existence" is discovered through broadcasts calls over the network and access is negotiated in a peer to peer relationship between the user and the service.

The next version of Microsoft's operating system will make considerable use of this new architecture model and in turn a new peer to peer authentication mechanism. Implications to current CGDI services, including WFS, are significant and are being addressed through proposals to incorporate security authentication within CGDI web services WMS and WFS. These changes are more than a year away from being included in the implementation specifications and correspondingly being endorsed for using within CGDI.

5.1.3 Digital Rights

Digital Rights is a concern with WFS since the data is passed to the end user and out of the control of the maintainer. Digital Rights are not a specific focus within WFS outside of the GetCapabilities operation within the metadata Service Identification section. The information within the Service Identification section is service level metadata which can include use restrictions and copyright under the ows:AccessConstraints element. This information does not address the potentially different copyright or use limitation information for each feature type or group served out within the same WFS response.

There are two options for WFS implementers if digital rights are to be incorporated into the current GML provided within a WFS response. The first would be to change the schema documents to provide custom elements in which digital rights information can be encoded. This would be a time consuming process and would require maintenance for each subsequent release of the GML standard.

The second option is to encode the digital rights information into an existing element present in all GML versions. Specifically the GMLBase element <description> could be used to not only describe the layer but also provide embedded digital rights information. This element is not the name of the GML layer and therefore would not interfere with application rendering of layer the resulting layer name. The <description> element exists in all GML versions and is readily available to users when looking and the properties of the layer. Generation of the information to fill this element could be a concatenation of a number of metadata fields ensuring that any digital rights information is maintained in one location but made available in the GetCapabilities response as well as the resulting GML.

The OGC is addressing digital rights management though a working group, GeoDRM WG, in which NRCan representatives participate. This working group's objectives are

- Enable business models for web-based geospatial services by identifying or developing a trusted infrastructure for purchasing and protecting rights to digital content.
- Guide the development of OGC specifications and best practices recommendations to permit the exploitation of mainstream DRM approaches, technologies and standards wherever possible
- Test, verify and mature as necessary the technologies required for geospatial DRM including electronic commerce and information security.
- Develop specifications for geospatial DRM that build on the OGC technical baseline.

The output standards of this group will define the service expectations with respect to digital rights.

5.2 Technical Implications

5.2.1 Standards Adoption

Currently the version of WFS utilized by most service providers is WFS 1.0.0. Implementation of WFS 1.1.0 by vendors has been extremely slow and this is reflected in adoption by service providers. The key indicator for adoption of a new WFS standard is its implementation within the University of Minnesota (UMN) MapServer open source project. MapServer has been widely used by CGDI GPs and with the WFS 1.1.0 capability not available in the current release, as referenced on the OGC compliance webpage, service providers cannot provide WFS 1.1.0 services without introducing new technology. With UMN MapServer not having implemented WFS 1.1.0 other vendors are not forced to implement WFS 1.1.0. Some vendors have implemented the new standard within their products and have highlighted the limitations of WFS 1.1.0.

With the recent release of GMLsf CGDI GPs are now faced with the having a GML specification that better suits their needs but is available through WFS 1.1.0 a standard their technology may not currently support. It will take the next year to see increased vendor adoption of WFS 1.1.0 using GMLsf. Once this occurs, specifically if it can be accomplished within UMN MapServer, then service providers will more readily implement WFS 1.1.0.

5.2.2 Performance

Related to Service Level Commitments, performance of WFSs is a key consideration. Performance is based on several factors as outlined in the section on WFS Process Performance (Section 3.4). WFS will impact WMS service capacity if the same server is used for both services. Server load limits are defined by maximum process size during peak requests. Current infrastructures of CGDI GPs have been calculated using WMS loads since WFS loads are very unpredictable, due to in consistent response GML file sizes.

WFS implementation would best be accomplished though a separate server dedicated to WFS service requests. The server could then be performance tuned to best suit the CGDI GP's individual data holdings. WFS will never have the response time of WMS even as file transfer rates over the open internet increase because the resulting files sizes required for WF will always be larger than the corresponding WMS.

The use of GMLsf is expected to improve performance of WFS 1.1.0 in comparison to the same data served using the GML 3 schema.

5.2.3 Distributed Data and Services

Currently the speed and simplicity of WMS allows data maintainers to also be the service providers of their data. Within the current services architecture of CGDI, applications access services closest to the source of the data. WMS and WFS implementations have the capability to cascade, meaning one service calls another service (Figure 4).

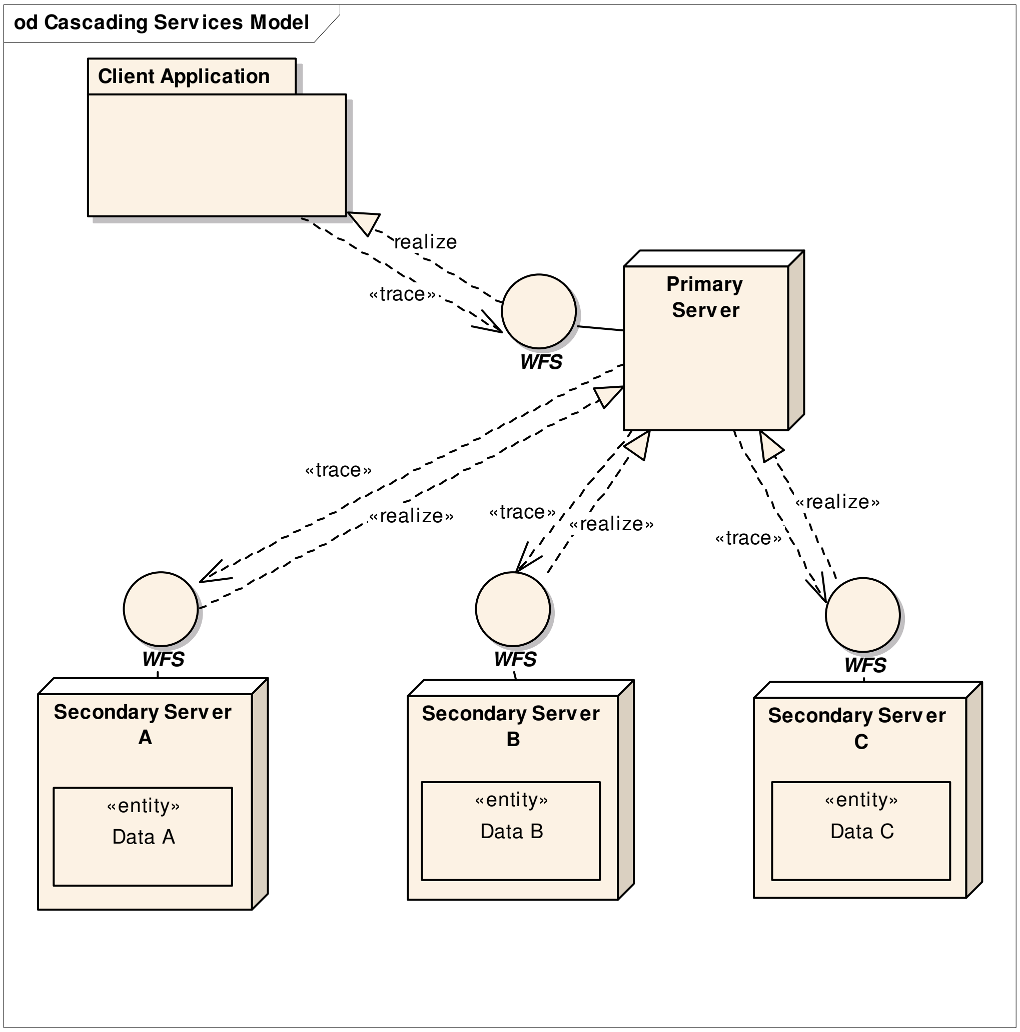

Figure 4: Cascading WFS

A variation on the cascading services model was developed as part of the OWS-3 test bed by OGC. For this test the WFS services were 1.0.0 compliant, contained different schemas and were dynamically transformed to produce a common WFS to a client application. The test implemented an “aggregation” service within the primary server that re-mapped each secondary server WFS schema to a single schema for return to the client application. By implementing a transformative service an additional process is added to the chain of services therefore adding to the time needed to return data to the client application. The benefit in this case would be that the data is in a defined schema regardless of the source WFS schema.

The alternative to cascading is always going directly back to source and having references to services within the primarily contacted service (Figure 5).

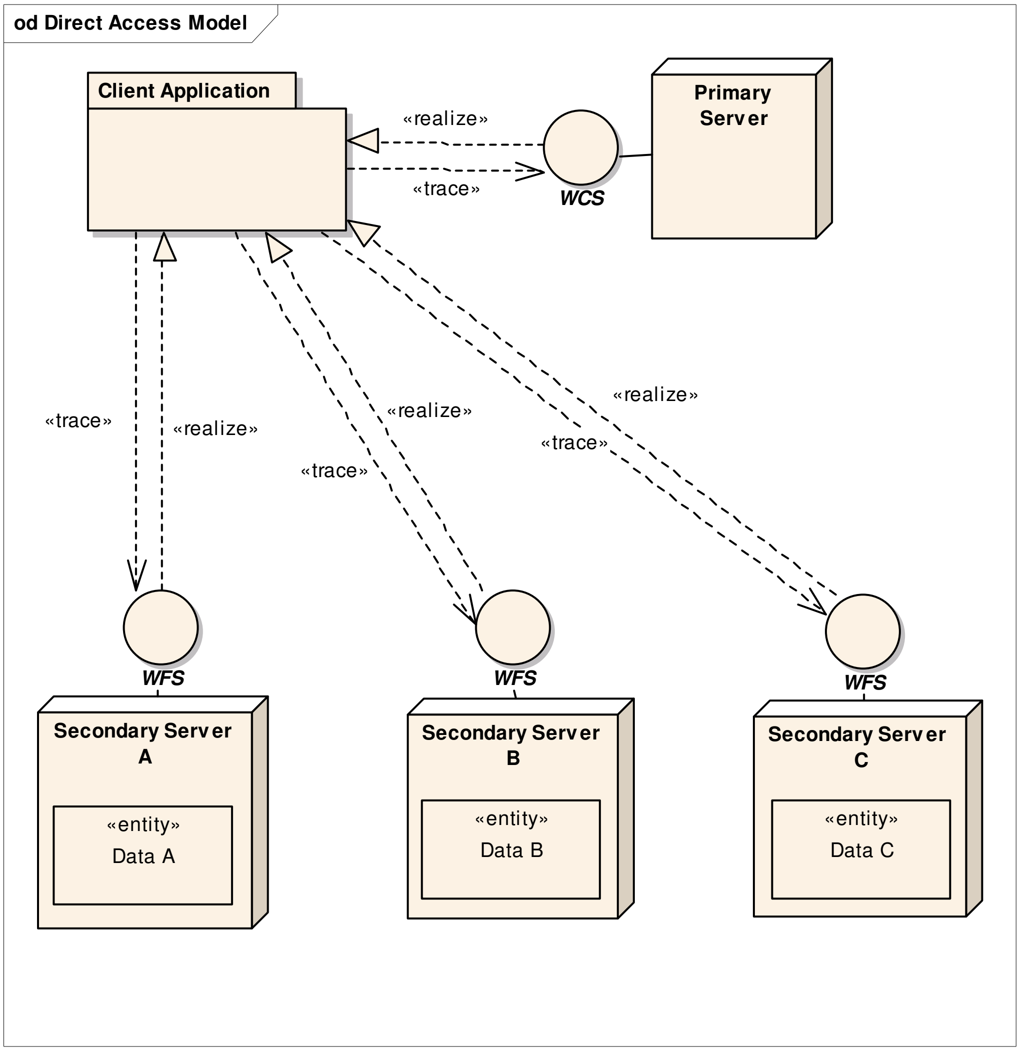

Figure 5: Direct Access WFS

There is a difference in performance between the cascading and direct models for WFS services. The cascading model requires an aggregation and response step that delays the final response to the client application. The cascading model requires additional resources in the primary server to accommodate the additional processing and the primary server becomes the single point of failure of the service request.

Direct access relieves the primary server of the processing requirement off loading any required processing to the client application. Individual services are therefore responsible for their own service availability and do not impact other services ability to respond. Under either model the potential to overload WFS services containing base map information is high.

A hybrid architecture for WFS called a “cascading cache” where the primary server would asynchronously update a dataset on the primary server (Figure 6). This would occur at off peak times and is not related to the WFS request from the client application. Data is updated from the source data provider service through the use of WFS. If necessary, data manipulations can be performed and the resultant data supplied as a new service, WFS or WMS, through the primary server.

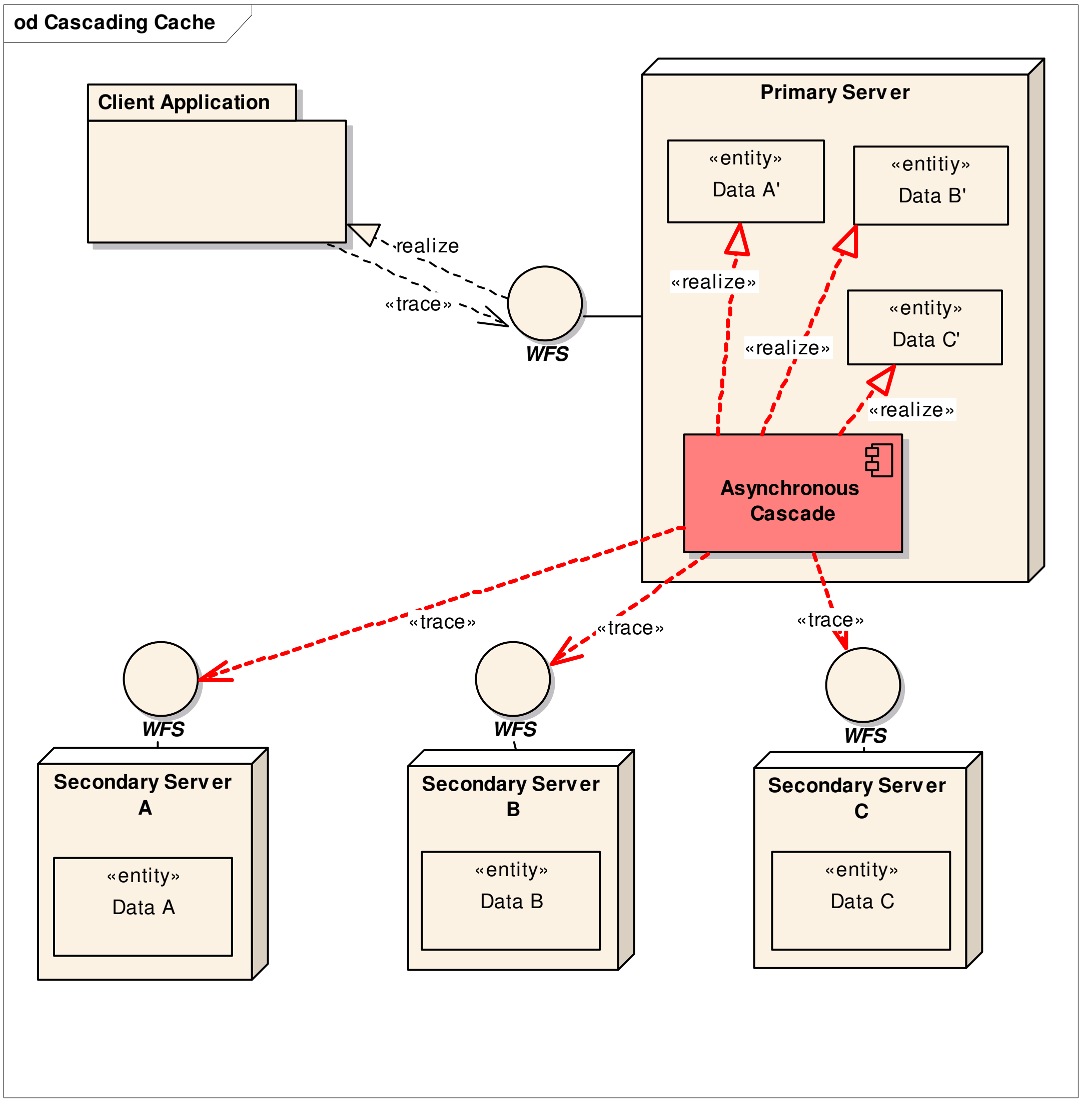

Figure 6: WFS Cascading Cache

The cascading cache model represents a potential change to the current update process such as those used for the GeoBase Portal. Asynchronous caching of data would provide the data access and update required within CGDI utilizing existing standards and specification and redistribute the server request loads to organizations focused on WFS serving of client applications. In addition the use of asynchronous updating of data can be limited to attribute data only allowing geometries and attributes to be managed by separate organizations and in turn served by a completely different organization.

5.2.3.1 Separation of geometry and attributes

For most data maintenance organizations attribute changes out number feature geometry changes. With the establishment of national frameworks, for example the National Road Network, organizations will maintain associated information while other organizations maintain the geometry.

WFS 1.1.0 XLink service class allows for the linking of geometry to externally referenced attributes. This uses the GML XLink:ref capability providing a discreet reference to an attribute value for each geometry attribute element. Implementing this level of linking adds tremendous processing time upon rendering of the WFS response GML. For each reference a call is made and a value is retrieved. Although not ideal for real-time applications the mechanism would be useful in maintenance operations where attributes need to be updated from a remote location associated to the geometries locally hosted (Figure 7).

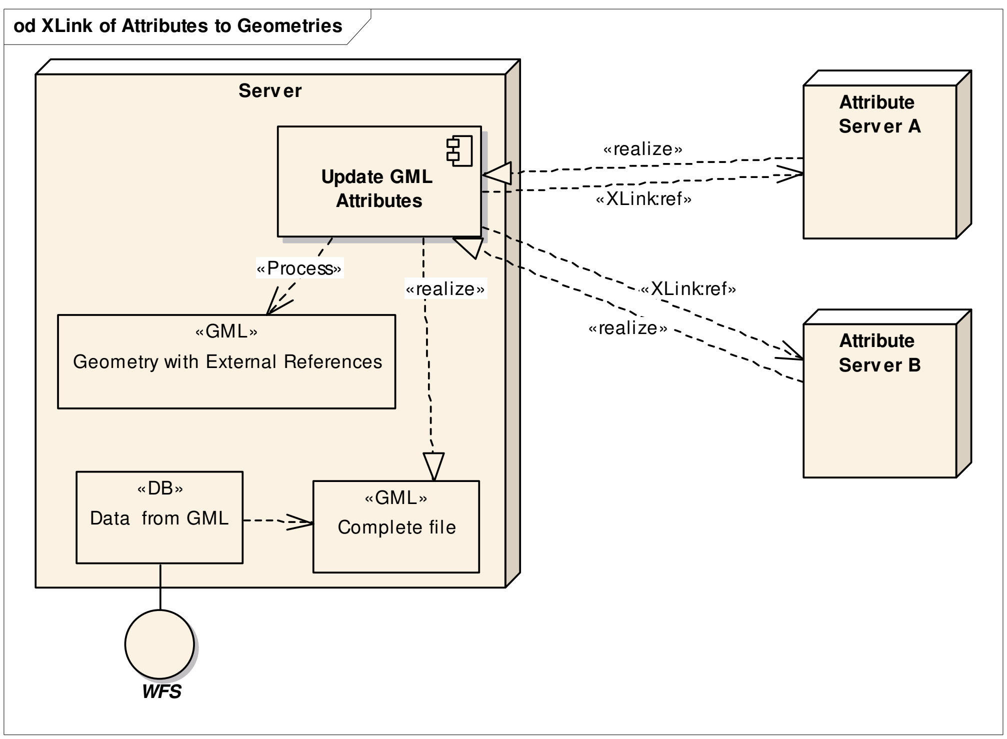

Figure 7: Asynchronous Attribute Update of GML

The association of geometries and attributes at runtime has been a key issue for some government programs. The National Land and Water Information Service (NLWIS) looked at this specific issue using WFS 1.0.0 and GML 2.1.2. NLWIS decided that the XLink service class did not provide adequate flexibility to implement their vision (Appendix A). Development of a framework specifically designed for manipulation and aggregation of disparate geometry and attributes has resulted in the OGC discussion papers on GeoData Access Services (GDAS) and GeoLinking (Appendix A).

5.2.4 Redundancy

As services face higher usage alternate sources of the same service need to be available to maintain business continuity of end users. This is applicable not only with government end users of services but increasingly CGDI services are the base for business decision support systems. When a service is overloaded or has a fault alternate identical services need to be available while the issue to the primary service is being dealt with.

The issue of redundancy at a service level is mirrored at the data level as well. A redundant service pointing to the same database is ineffective if there is a database fault. The notion of replication of data may be considered contrary to the services model within CGDI. Although CGDI services are not “mission critical” the increasing reliance on CGDI services will necessitate a need for redundancy and therefore replication of service interfaces and associated data in multiple locations within the CGDI.

5.2.5 End User Access

Currently users have client applications that are not WFS 1.1.0 compliant. In addition client applications do not read GML produced by different vendors consistently. Client applications will go through considerable enhancement within the next 2 years to accommodate WFS 1.1.0 and GMLsf as the new benchmark for vector based interoperability.

Also during that time the movement to the peer-to-peer services architecture will require users to either use existing CGDI services, like WFS, as is or move to the new service security architecture. During this transition CGDI will need to adapt to changing user needs and at the same time indicate to vendors the new security services architecture of CGDI. Only in a lead position can CGDI facilitate vendor implementation of any new services architecture and therefore meet users’ expectations.

6. Recommendations

With very few organizations implementing WFS 1.1.0 and with GMLsf just surfacing as an implementation specification it is difficult to advocate a single course of action that would be equally effective to all CGDI GP's.

However, there are some recommendations focused on key areas:

1. WFS should be implemented by organizations that maintain vector data;
2. WFS services should be initially used as data process and access interfaces and not as client application interfaces;
3. Investigate the detailed impacts of a new services security architecture based on the peer-to-peer model and certificate authorization;
4. Implement a WFS cascading cache capability to address redundancy, performance and service level commitments; and
5. Implement digital rights text within the WFS response GML.

Although some of these recommendations can be implemented immediately without broad impact others require a more comprehensive understanding of the near future, 1-3 years, operational architecture of CGDI.

Looking forward there are two primary steps that can be taken to ensure proper adoption and effective use of WFS within CGDI. The first is to develop a detailed operational architecture and the second is for CGDI/GeoConnections to support implementers of WFS with more compliancy tools and guidelines.

6.1 Detailed Operational Architecture

CGDI GP's require context for the implementation of WFS. Communication of the larger picture with sufficient detail will allow partner organizations to identify their roles and define the service implications relative to their own service infrastructures. The most effective way to accomplish this is to develop a detailed operational architecture profiled specifically to use of vector data, WFS and GML. This detailed operational architecture would provide not only an implementation road map for WFS and GML but also address issues of security architectures and digital rights management.

The first step in developing an operational architecture is to have an understanding of the discreet business and system activities related to serving data using WFS. This would include not only identifying activities within the system but also interactions with business and system activities outside of the CGDI WFS boundary. This information is usually captured using UML, a modeling technique to diagram processes and structures for business and system activities. UML was used for diagrams within this document.

The UML diagram and associated textual content are defined within a Use Case Functional Specification (UCFS). The resulting UCFS would define the operational architecture for WFS within CGDI in enough detail to allow individual partners to define how they are going to contribute to the larger WFS vision.

6.2 CGDI/GeoConnections

To better allow CGDI GP's to implement WFS into its service architecture the following CGDI structures would enable the resulting services to be consistent with other CGDI implementations of WFS and confirm compliance with the operational architecture.

6.2.1 Technical Capabilities Review

Test beds for WFS clients, WFS GML streams, GMLsf and XLink would allow CGDI data providers as well as application developers to achieve a functionally interoperable capability using WFS/GML.

CGDI has shied away from endorsing specific technologies regardless of their adherence to its standards and specifications and as a result government departments have had difficulty in determining which technologies are truly interoperable. With standard interfaces now moving beyond WMS there is a need for the CGDI to provide mechanisms to compliancy test software within the context of CGDI and therefore transparently identify solutions that provide best value to CGDI participants and users.

6.2.2 Guidelines for WFS

Although GeoConnections II now focuses on users' needs there is still a need to provide support to government departments implementing WFS. This support would be geared towards help not on the technical implementation equally focused on the business decision process related to releasing vector data outside of an organization. A WFS Guidelines program would allow for the definition of best practices for the implementation of WFS. Specifically it would provide participants in CGDI the checklists and business decision trees for implementing, configuring and participating in CGDI.

Previously CGDI has not wanted to impose process on CGDI participants and the result has been that government departments are struggling with the business implications, data rights and service liability issues, as much as with the technical implementations of WFS. By providing case studies on how these issues have been addressed adoption of WFS and GML datasets within CGDI will become easier and users will have a clearer understanding of the overall operational considerations of WFS.

Appendix A: National Land and Water Information System

The National Land and Water Information System (NLWIS) has been a leader in the development of interoperability standards and working with GeoConnections to advance CGDI capabilities through interoperability prototypes.

NLWIS utilizes WFS as a service interface but has invested time and effort in advancing the linking of features and attributes allowing for single geometries to have multiple sets of associated attributes. Each attribute set could be maintained by a different group but the common geography allows users to contrast, compare and analyse attributes in a spatial context. Figure 5 is a simple interaction diagram showing the requests and processes involved in linking disparate geographic features and attributes to produce a single layer within a map. The output for the map layer can be WMS, WFS or GML.

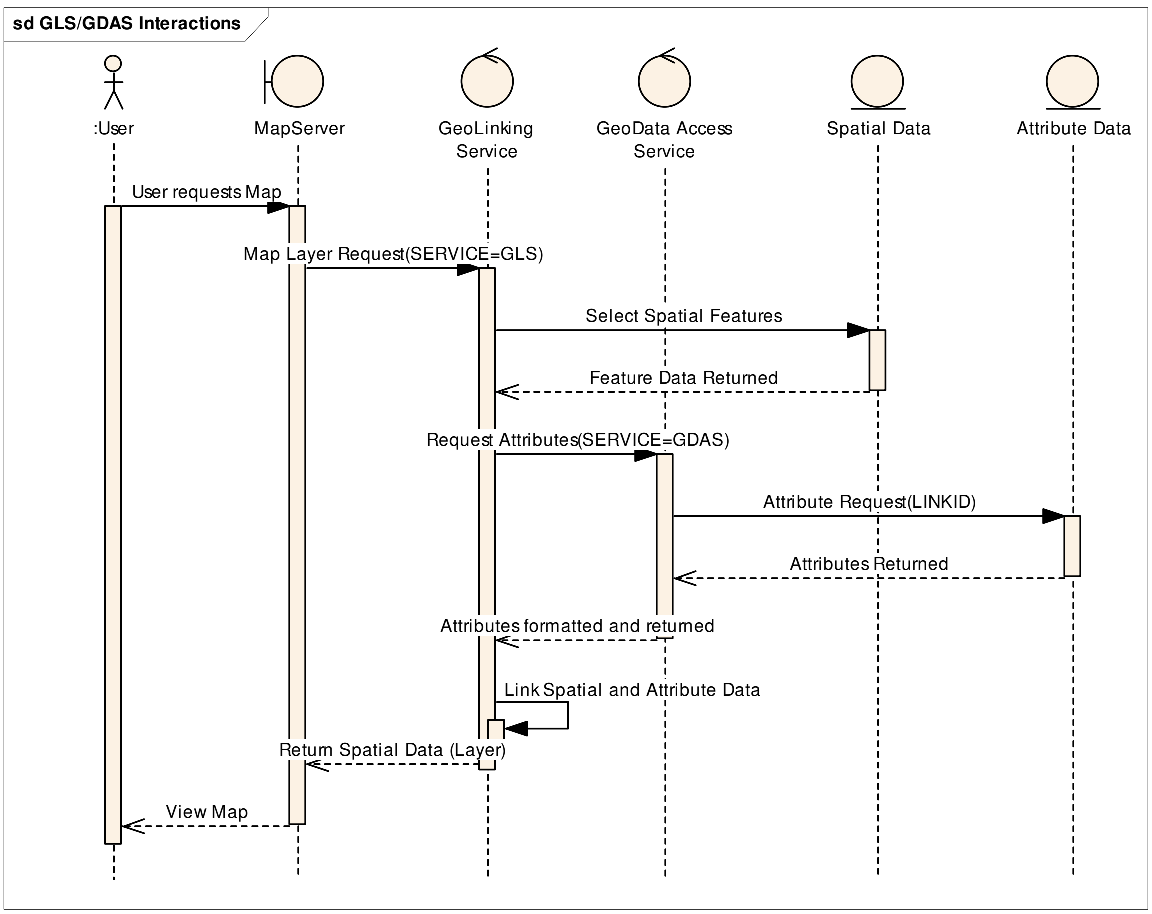

Figure 8: GeoLinking and GeoData Access Service interaction diagram

NLWIS did consider the use of WFS 1.1.0 and GML 3 capabilities for XLink before the standards were released. NLWIS investigated the feasibility of constructing a GML schema that would allow the linking necessary for NLWIS. The complexity of the resulting GML schema prompted further investigations into GDAS and GeoLinking as potential standards to provide a simpler and flexible capability to reference features and attributes at runtime. GeoLinking is a primary mechanism to serve linked features and attribute datasets. The GeoLinking XML coding provides the same functionality as the XLink concept that was implemented in GML 3, but is much less verbose. In addition the path NLWIS needed to take in allowing non-spatially referenced attribute data to be processed meant that WFS and GML still did not have the capabilities required for NLWIS's increasing complex web processing requirements.

