Validating Web Feature Server
Final Report

Submitted To: Program Manager
GeoConnections
Victoria, BC, Canada

Submitted By: Jody Garnett
Refractions Research Inc.
Suite 400 - 1207 Douglas Street
Victoria, BC, V8W-2E7
jgarnett@refractions.net
Phone: (250) 383-3022
Fax: (250) 383-2140

TABLE OF CONTENTS

1 EXECUTIVE SUMMARY

The goal of the Validating Web Feature Server project was to create a spatial validation framework that could be reused in enterprise applications, and also a working integration of that framework within an existing open source Web Feature Server. The project also included enhancements to the underlying Web Feature Server as a goal – the enhancements included support for proprietary spatial databases, and a web-based configuration system.

The project partners included Refractions Research, the Open Planning Project (TOPP), the BC Ministry of Sustainable Resource Management (MSRM), and GeoInnovations. Refractions, TOPP, and MSRM provided programming and spatial domain expertise as contributions in kind, and GeoInnovations provided matching funds.

The results of the project were positive: our project delivered the Validation framework, and also integrated it into the GeoServer Web Feature Server. The Validation framework and the GeoServer WFS are both released as "open source" software, which means that other software developers can look at the internals, make changes if they wish, and redistribute the software as they wish.

Our company is now very experienced with Web Feature Server technology, and with the GeoServer Web Feature Server internals in particular. We are using that expertise in our consulting business, and will be using it in sales activities over the coming year.

The products of our project provide technical benefits to many CGDI members:

- for members without a Web Feature Server, the Validation framework can be used to construct data testing and validation engines than can be run on database and file based spatial data;
- for members using a Web Feature Server, the Validation engine can be integrated directly in to the editing workflow, and protect the integrity of the underlying data on the fly;
- for members with proprietary database servers, such as OracleSpatial and ArcSDE, the project added the ability to expose those databases using the GeoServer WFS; and,
- for all members with spatial databases, the project included a substantial performance and integrity upgrade to the locking facilities in the GeoServer Web Feature Server.

In general, the entire technical community within CGDI can benefit, because the results of the project have been released as open source software. We feel this project was a positive for ourselves, and for the CGDI community, and look forward to seeing our work deployed by CGDI network members.

2 INTRODUCTION

The goal of this project is to build a Validating Web Feature Server (VWFS) by adding extra functionality to a Web Feature Server (WFS). A Web Feature Server is an application that delivers information about spatial objects across the Web.

The WFS that we have used is called GeoServer. GeoServer is an open source implementation of the Open GIS Consortium's Web Feature Server Specification. Our project will add a level of validation to GeoServer to help maintain spatial database integrity.

3 PROJECT OUTCOMES

This project has resulted in the release of GeoServer 1.2.0, making use of the GeoTools 2.0 B2 library.

As a result of this projects contributions GeoServer 1.2.0:

- is compliant with the Open GIS WFS specification;
- contains an integrated Validation Processor;
- allows for configuration via a web based user interface;
- allows strong transaction support;
- provides in process transaction support for shapefiles; and
- is substantively faster then previous releases

As a result of this project's contributions GeoTools 2.0 B2 includes:

- an advanced DataStore API with strong transaction support and light weight transaction and locking support for file based spatial data
- a Validation Process and Validation Framework with 42 Validation Plug-Ins covering a range of spatial validation concerns
- a graph module allowing networks to be created from spatial information and advanced validation tests to be performed
- Improved support for Postgis, Oracle and ArcSDE

3.1 Comparison to Objectives

The objective of this project is to create and bundle a Validating Web Feature Server (VWFS) based on the Java language and executable within a J2EE infrastructure.

Project objectives:

- be fully compliant with the OpenGIS WFS specification;
- inter-operate with existing commercial and open-source spatial databases;
- support WFS update validation based on attribute and spatial restrictions;
- support WFS update validation based on plugable validation modules;
- be written in platform independent Java for a J2EE environment; and
- support using WFS as a feature source (i.e., cascading WFS).

This project has succeeded in meeting our main objectives:

- has passed OGC compliance testing;
- provides support for existing spatial databases and file formats;
- is a J2EE based application;
- provides plug-in based validation test suites; and
- includes 42 plug-ins providing both attribute and spatial restrictions

This project has not created a “Cascading WFS”. The use of GeoServer as a cascading WFS was not a priority with our project partners at the British Columbia Ministry of Sustainable Resource Management (MSRM). MSRM advocated directing our energies at providing robust support for Oracle and ArcSDE.

3.2 Technical Discussion

3.2.1 Validating Web Feature Server

The creation of a Validating Web Feature Server is the main objective of this project and represents a technical success in its own right.

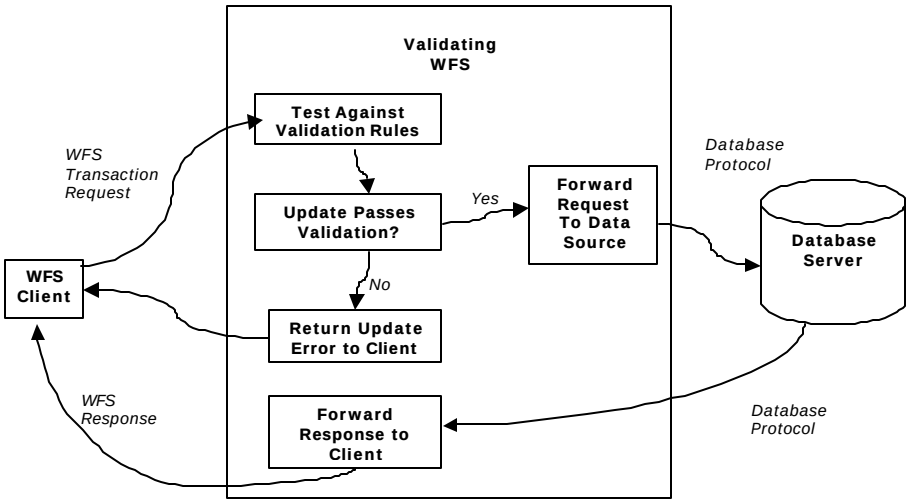

Figure 1: Validating Web Feature Server Workflow

The project has added a Validation Processor to the GeoServer J2EE Web application. This smoothly integrates the validation process in the Web Feature Server workflow.

The Validation Processor is used as a service by the WFS Transaction Operation.

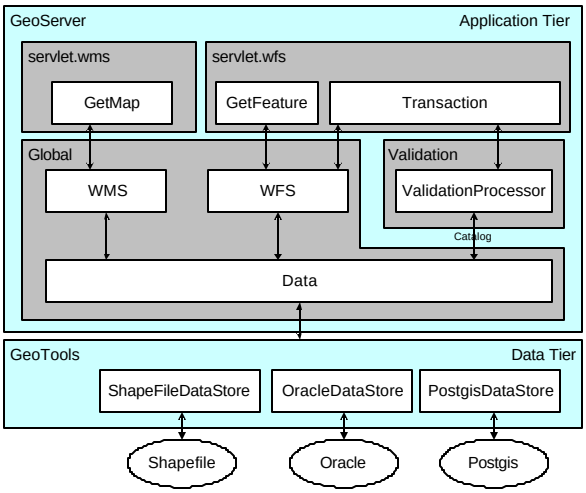

Figure 2: GeoServer Layer Diagram with Validation Processor

The validation processor is called on to perform validation at three stages of the Transaction Operation execution:

1. During the processing of Inserts the Validation Processor is used to ensure that the Features being added are internally consistent.
2. During the processing of Updates the Validation Processor is used to ensure that each Feature is internally consistent after the update has been performed
3. After all Update, Insert and Delete operations have been processed the Validation Processor is used to check the integrity of the Spatial Database(s).

3.2.2 Web Based Configuration

This project has extended the GeoServer with a user interface for configuration. This has been accomplished without changing the configuration file format from previous GeoServer releases.

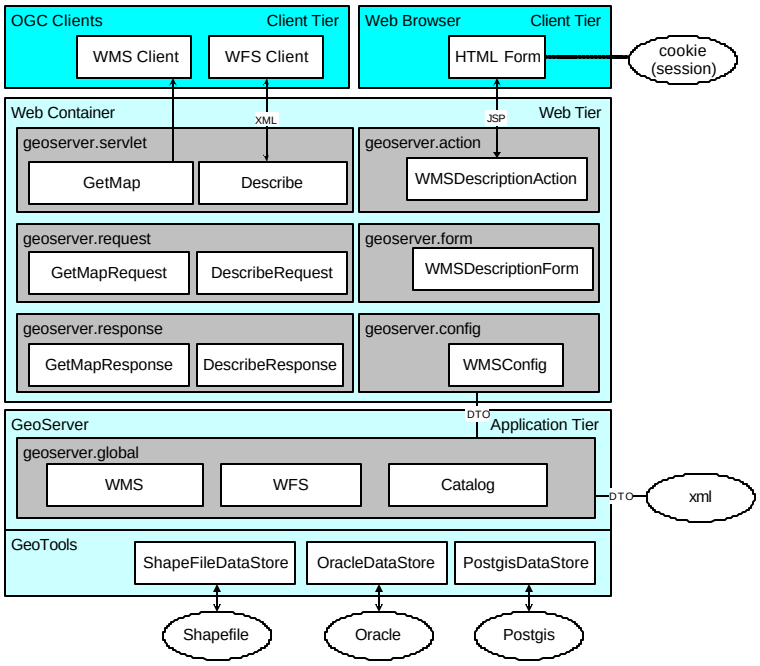

Figure 3: GeoServer Configuration Layer Diagram

The design for the GeoServer Configuration System has succeeded in:

- Providing GeoServer Web Based Configuration
- Providing dynamic feedback from the GeoServer application
- Providing a consistent user experience
- Provided a robust configuration framework for GeoServer, capable of independent unit testing
- Allowing for the smooth integration of new options such as the configuration of WFS level of service

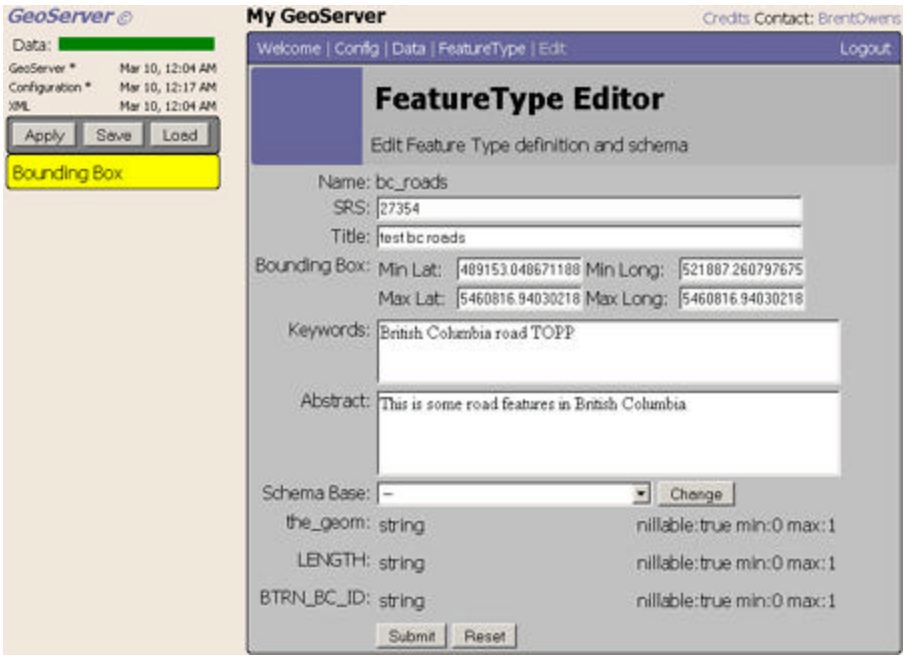

Figure 4: Feature Type Editor

A web based configuration system represents the most requested feature for GeoServer, and we are pleased to have been central to its introduction.

The web based configuration system significantly reduces the setup cost associated with GeoServer.

3.2.3 DataStore API and Strong Transaction Support

This project provided impetus to the GeoTools2 data source revision resulting in the GeoTools DataStore API. Through our close involvement the DataStore API it has been strongly aligned with the goals of this project and the needs of GeoServer.

The GeoTools DataStore API:

- Is modeled after accessing a "Stream" of Spatial Data
- Implements Transactions separately from a DataStore
- Implements Feature Locking as part of the API (not as an extension)

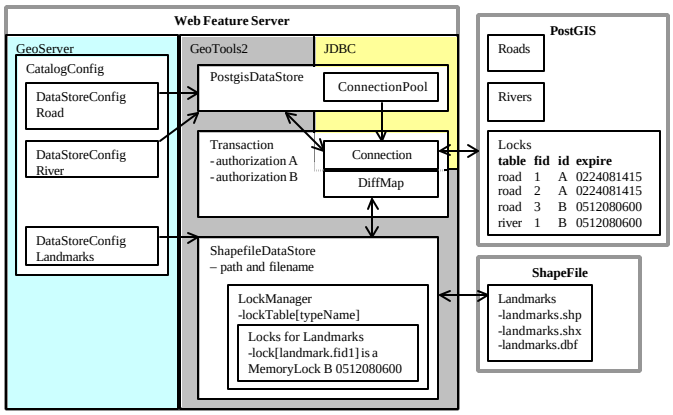

Figure 5: GeoServer WMS Branch and New GeoTools DataStore

The DataStore API has several benefits for GeoServer:

- Transaction allows shared use database connection
- Strong Transaction Support is provided
- Transaction and Locking are provided for file DataStore
- Scalability is not limited by memory

The DataStore API provides a significant performance improvement over the original GeoTools "DataSource" API. The following graph indicates the time, in seconds, to read a dataset vs. the size of the dataset.

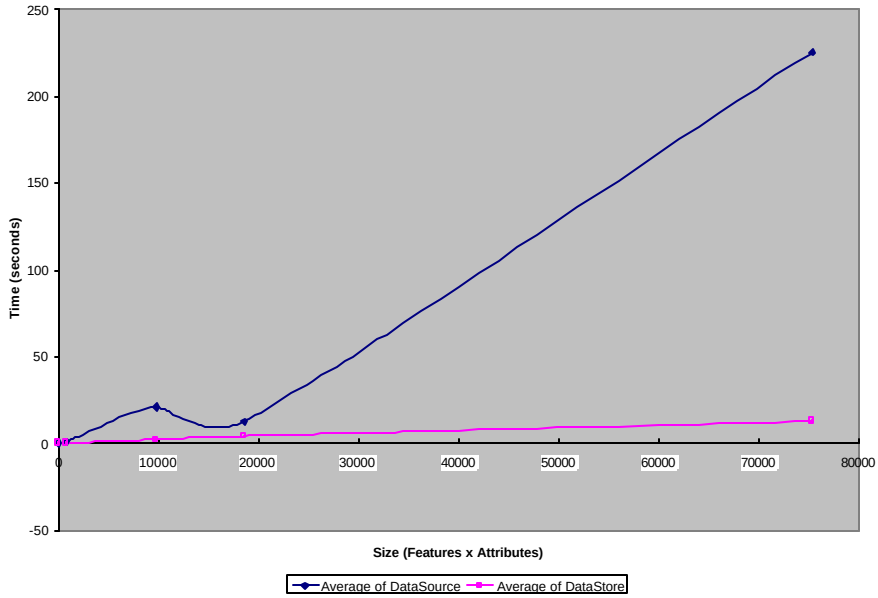

Figure 6: DataStore / DataSource Performance Comparison

Size is measured as the number of features in the dataset multiplied by the number of attributes associated with each feature.

3.2.4 Validation Processor and Framework

This project has developed a Java based Validation Framework and Validation Processor.

The Validation Framework provides a strong Plug-In Architecture that has matured through several iterations of development and feedback.

Validation API:

```
public interface Validation {
    void setName(String name);
    String getName();
    void setDescription(String description);
    String getDescription();
    int getPriority();
    String[] getTypeRefs();
}
public interface FeatureValidation extends Validation {
    boolean validate(Feature feature, FeatureType type, ValidationResults results)
        throws Exception;
}
public interface IntegrityValidation extends Validation {
    boolean validate(Map layers, Envelope envelope, ValidationResults results)
        throws Exception;
}
```

The validation API provides two interfaces for Plug-Ins:

- FeatureValidation used for testing of a single Feature for internal consistency
- IntegrityValidation used for testing Integrity of one or more Feature Types

The Validation Framework features a well-defined validation language for the definition of test suites.

Sample Validation test suite from GeoServer 1.2.0 installation:

```
<?xml version="1.0" encoding="UTF-8"?>
<suite xmlns="testSuiteSchema"
       xmlns:gml="http://www.opengis.net/gml"
       xmlns:ogc="http://www.opengis.net/ogc"
       xmlns:xs="http://www.w3.org/2001/XMLSchema-instance"
       xs:schemaLocation="testSuiteSchema /data/capabilities/validate/testSuiteSchema.xsd">
       <name>IntegrityCheck</name>
       <description>Integrity Check Description</description>
       <test>
               <name>FID Uniqueness</name>
               <description>Checks all Features for ID Uniqueness</description>
               <plugin>FID Uniqueness</plugin>
               <argument >
                       <name>attributeName</name>
                       <string>id</string>
               </argument>
               <argument >
                       <name>typeRef</name>
                       <string>*</string>
               </argument>
       </test>
       <test>
               <name>Is Valid Geometry</name>
               <description>Is Valid Geometry</description>
               <plugin>Is Valid Geometry</plugin>
               <argument >
                       <name>typeRef</name>
                       <string>*</string>
               </argument>
       </test>
</suite>
```

The Validation Processor and Framework may be used independently of the GeoServer application. A batch tool has been provided that makes use of the Validation Processor.

3.2.5 Validation Plug-Ins

Over the course of this project, validation plug-ins have been produced in a number of different categories.

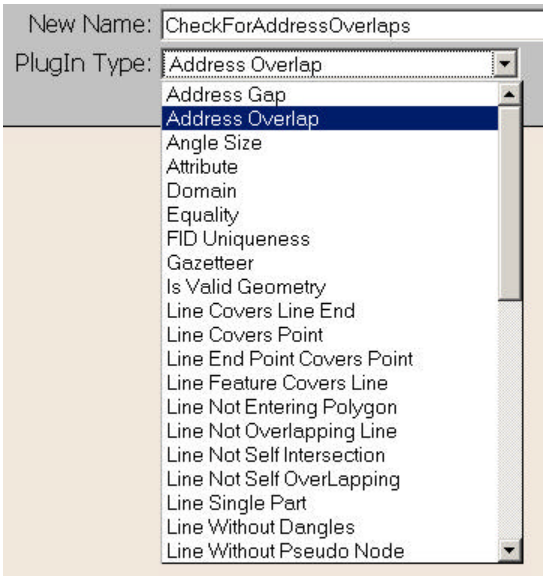

Figure 7: Selecting a Plug-In using GeoServer 1.2.0

Highlights of this development process have included:

- Porting a well-documented Network Builder from the JUMP Feature Model to the Geotools2 Feature Model allowing the Validation Processor access to powerful technology from another sphere of open source development.
- The creation of a Gazetteer-based plug-in illustrating the use of multiple OGC-based services in a single workflow
- A total of 42 Plug-Ins have been produced

4 BENEFITS REALIZED

4.1 Benefits to Canadians

The Validation Framework and the enhanced GeoServer capabilities delivered by this project will allow Canadian government agencies to maintain data quality and move to standards-based data access without purchasing expensive proprietary software. They will help protect the quality of Canadian data as the CGDI moves to a more and more distributed data management model. And they will provide smaller organizations (NGOs, municipalities, etc) the opportunity to become involved in the CGDI network without substantial spending on GIS server software.

4.2 Benefits to the Project Team

4.2.1 Refractions Research

This project has been an excellent opportunity to undertake a pure research and development exercise, of substantial scope, and to study problems that are within our realm of professional expertise. Through this project, we have gained substantial expertise in the GeoServer and GeoTools open source code bases.

4.2.2 The Open Planning Project

The VWFS Project has been of significant benefit to The Open Planning Project(TOPP). The project has focused attention on TOPP's GeoServer project, and provided a pool of active developers working on GeoServer.

4.2.3 British Columbia Ministry of Sustainable Resource Management

This project has provided MSRM with several key benefits:

- It provides a data distribution and maintenance tool which can be used at all levels of government and in the private sector, regardless of budget
- Provides a Validation Framework
- Lowers the cost of maintain GeoServer as a WFS through improved configuration and the addition of a user interface

4.3 Benefits to Open Source Projects

The Project Team has synchronized its development efforts with two dynamic open source communities: GeoServer and GeoTools.

4.3.1 GeoServer

The benefits released by the GeoServer project have been substantial both in terms of documentation and code contributions:

- Validation as a new capability of their Transaction operation
- Web Based Configuration
- Substantive development support and documentation

4.3.2 GeoTools

GeoTools has received several benefits including:

- The VWFS project directly influenced the development of the GeoTools DataStore API. The emphasis this project has placed on performance and strong transaction support has been realized in the DataStore API.
- The Validation Framework is being hosted as a GeoTools module.
- A Graph module is being hosted as a GeoTools module

This project has also resulted in the addition of Refractions staff to the GeoTools Project Management Committee.

5 RELATED COMMERCIAL ACTIVITIES

Refractions Research is an open source spatial services company, and requires expert knowledge and qualifications to backstop our value as a consultancy to agencies implementing open source solutions. In addition to producing a useful product and programming tools, this project has added to our corporate store of knowledge.

The products of this project (Validation Framework, WFS integration) are part of our marketing efforts, to bring open source spatial solutions to governments and organizations across North America.