Geophysical Research Abstracts
Vol. 13, EGU2011-10534, 2011
EGU General Assembly 2011

Enabling collaborative editing with synchronization of distributed geospatial databases

David Arctur (1), Panagiotis (Peter) Vretanos (2), and Jeff Harrison (3)

(1) Open Geospatial Consortium, Interoperability Program, Austin, United States (darctur@opengeospatial.org), (2) CubeWerx Inc., Toronto, Canada (pvretano@cubewerx.com), (3) The Carbon Project, CubeWerx USA, Washington DC, United States (jharrison@thecarbonproject.com)

Agile spatial data infrastructures, based on OGC standards, allow data provider organizations to deploy increasingly efficient networks capable of responding to dynamic requirements for using geospatial data. In this context, data provider organizations are being more and more called upon to deliver current, timely and verified data over the World Wide Web.

In order to satisfy these requirements, data providers must collaborate with outside entities to collect new data and/or update their existing data holdings. This could mean synchronizing their data with closest-to-source providers, such as between municipal, state/provincial and federal levels of government. This could also mean crowd-sourcing their data production or supporting volunteer geographic information (VGI). Crowd-sourcing or VGI has proven to be a very effective tool for emergency management / disaster relief situations as was shown by the response to the Haiti earthquake of 12-January-2010.

Regardless of the nature of the collaboration, there is a need for a service to mediate the interaction between data providers and outside entities acting as data collectors. This paper describes a Geosynchronization service (candidate standard) which would be deployed by a data provider, that sits between the provider's data store(s) and any number of data collectors. It allows data collectors to submit new data or make modifications to existing data without directly affecting the data in the provider's data store(s) until validation has been applied thus ensuring that the data published by the provider is of high quality.

This standard makes no assumption about the nature of validation or quality criteria a provider may have established; only that the Geosynchronization service allows the data to be evaluated for quality before being applied to the provider's data store(s). For example, changes submitted by trusted data collectors may be applied directly to the provider's data stores while data obtained from external sources would be subjected to more rigorous validation.

Interested parties (or subscribers) will be able to flexibly track which changes have been submitted by data collectors, what the disposition of those changes was (i.e. whether the changes were accepted or rejected by the validation process) and also track the actual changes made to a provider's features for the purpose of replication. This flexible notification capability is achieved by allowing subscribers to specify predicates that precisely identify about which events they are interested in being notified. For example, a subscriber might register that they are interested in receiving notifications about proposed changes made within a specific geographic area.

Finally, this Geosynchronization service will support replication of the data provider's geospatial features by maintaining an event channel with a log of all changes made to the features. This channel may be used to support several modes of replication including having the Geosynchronization service read the change log and apply the changes to one or more target feature types that have been subscribed to for replication. Replication may be full or partial, possibly involving a schema translation from the provider's feature schema to the schema of the target feature.

This paper presents the current design of the OGC Geosynchronization services, with intended applications for a number of different scenarios, as well as lessons learned and next steps.