

GEOPORTAL CLOUD

Roman K. Fedorov, Alexander S. Shumilov, Gennady M. Ruzhnikov

Matrosov Institute for System Dynamics and Control Theory of Siberian Branch of Russian
Academy of Sciences, Irkutsk, Russia

Abstract

Within the article, a technology is proposed for creating a geoportal for the collection and exchange of data between subject specialists. The technology is based on the developed model geoportal that implements common functions for many information systems. The technology significantly accelerates the creation of a subject-oriented geoportal that implements a ready-made set of various functions.

Keywords: Geoportal, services, WMS, WPS, spatial data

ОБЛАКО ГЕОПОРТАЛОВ

Фёдоров Р.К., Шумилов А.С., Ружников Г.М.

Институт динамики систем и теории управления СО РАН, Иркутск

В рамках статьи рассматривается технология создания геопортала для сбора и обмена данными между предметными специалистами. Технология базируется на разработанном типовом геопортале, реализующим общие для многих информационных систем функции. Предлагаемая технология значительно ускоряет создание предметно-ориентированного геопортала, реализующего готовый набор различных функций.

Ключевые слова: геопортал, сервисы, WMS, WPS, пространственные данные.

Введение. Развитие сетей передачи данных привело к активному развитию и внедрению Интернет технологий. Переводятся для работы в браузере настольные приложения (Adobe Photoshop, Microsoft Office и т.д.). Созданы и активно используются облачные хранилища данных, такие как Google Drive, Dropbox, Яндекс.Диск и др. Развиваются облачные технологии, позволяющие удаленно работать со своими данными и приложениями в любой точке мира, где имеется Интернет.

В настоящее время существует тенденция к интеграции, т.е. обмену данными и методиками анализа/обработки, на основе современных технологий. В основной акцент делается на интеграцию данных. Разрабатываются информационные системы (ИС), с помощью которых специалисты обмениваются информацией. Отметим ряд подобных проектов: орнитологический ресурс eBird [1]; широкую известность среди ботаников России получил использующий краудсорсинг открытый атлас и онлайн-определитель растений России и сопредельных стран Плантариум [2]. Ныне действует ряд глобальных информационных ресурсов, ориентированных на поддержку научных исследований биоразнообразия: Global Biodiversity Information Facility [3], Catalogue of Life [4]; фиторазнообразия: The Plant List [5], The International Plant Names Index (IPNI) [6] и др. В России существуют информационные системы ООПТ России [7, 8], Зоологического института РАН [9], CRIS (Cryptogamic Russian Information System) [10] и др. Перечисленные ИС ориентированы на ввод и редактирование отдельных, заранее определенных таблиц с заданной структурой данных. Добавление новых видов объектов, отличающихся атрибутивным составом, классификаторами, потребует значительной модернизации этих систем. В настоящее время, активно развиваются и внедряются международные стандарты представления, обмена и сервисов обработки пространственных данных через Интернет. Однако перечисленные ИС не поддерживают передачу данных в соответствии с этими стандартами и их обработку различными сервисами. Комплексирование данных перечисленных систем является нетривиальной задачей из-за отличия структуры, использования разных справочников, программных интерфейсов и т.д. Создание новой системы на текущий момент занимает значительное время и часто требует программирования.

Основная идея. Большинство ИС, ориентированных на поддержку научных междисциплинарных исследований, имеют аналогичные функции и различаются в основном структурой используемых данных. В ИДСТУ СО РАН реализована технология создания ИС в облачной среде. Базируется технология на разработанном типовом геопортале, реализующим общие для систем функции. Некоторые из функций вынесены на отдельные узлы облачной среды.

Общая архитектура облачной среды. Архитектура облачной среды представлена на рис. 1. Пользователь работает с помощью браузера.

Типовой геопортал разработан на основе программной платформы Node.js и СМС Calipso.

«Каталог данных» — предназначен для хранения метаданных таблиц и их структурных спецификаций. Каталог обеспечивает поиск данных по всем геопорталам.

«Сервисы» — вычислительные узлы виртуальной среды, выполняющие методы обработки данных. Методы должны быть реализованы в соответствии со стандартом WPS.

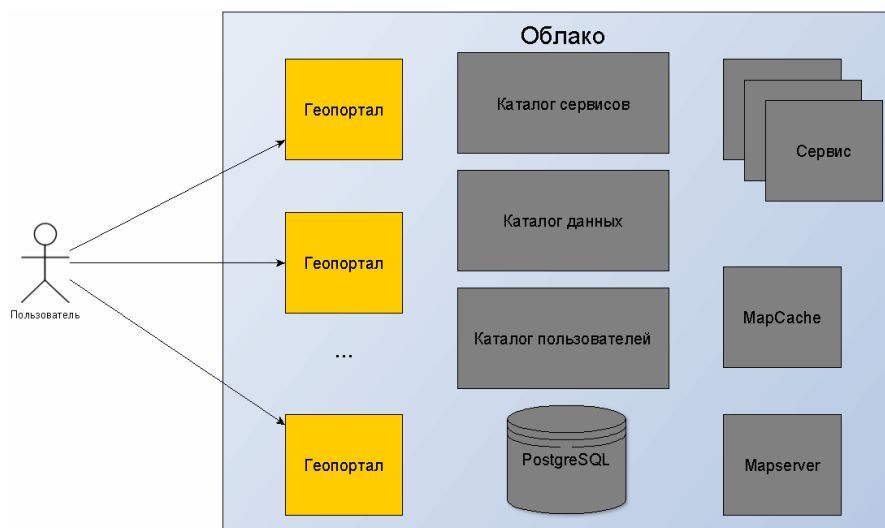


Рис. 1. Архитектура облачной среды.

«Каталог сервисов» — предназначен для хранения метаданных сервисов. Каталог обеспечивает поиск сервисов и их выполнение.

«Mapserver» [11] — проводит на стороне сервера генерацию изображений слоев карт в соответствии со стандартом WMS. Выделен на отдельный вычислительный узел из-за необходимости использования 80 порта. Многие организации блокируют доступ к остальным портам;

«MapCache» [12] — проводит кэширование карт для получения скоростного доступа.

«PostgreSQL» — СУБД [13]. Для обработки пространственных данных в СУБД установлено расширение PostGIS. Каждому пользователю системы предоставляется схема базы данных, в которой он может создавать таблицы. СУБД может функционировать на каждой виртуальной машине геопорталов или на общей СУБД.

Все рассмотренные компоненты архитектуры функционируют в рамках облачной среды. Разработанная технология базируется на рассмотренной архитектуре и состоит из следующих этапов:

- 1) развертывание нового узла облачной среды на основе типового геопортала и присвоение DNS имени, что позволяет получить общедоступный сайт в сети Интернет;
- 2) разработка таблиц для хранения данных пользователей на основе структурных спецификаций;
- 3) модификация пользовательского интерфейса.

Рассмотрим подробнее все этапы предлагаемой технологии.

Развертывание геопортала в облачной среде. Виртуальные машины создаются на основе шаблонов (в терминологии VMware это "template", в Openstack это "image"), где заранее установлен и сконфигурирован типовой геопортал. Типовой геопортал содержит программную систему ввода и редактирования данных, которая обеспечивает: многопользовательскую работу через Интернет, ввод и редактирование данных разных по структуре реляционных таблиц и т.д.

Геопортал содержит средства контроля и модерации ввода данных. Пользователь, создавший таблицу, автоматически назначается её владельцем. Для каждого пользователя владелец может установить права доступа:

- Редактирование собственных записей. В этом случае пользователь может создавать, редактировать и просматривать свои записи в таблице.
- Просмотр записей других пользователей.
- Редактирование записей других пользователей.
- Модератор. Пользователю разрешается редактировать записи других пользователей и устанавливать признак, что запись прошла модерацию.

Если владелец не установил «Просмотр и редактирование только владельцем», то таблица отображается в каталоге таблиц. Любой зарегистрированный пользователь может отправить запрос владельцу на получение дополнительных прав.

Геопортал поддерживает регламентированную передачу данных WPS сервисам [14]. При запуске WPS сервиса пользователь может выбрать таблицу в качестве входного параметра. Перед совершением непосредственного запроса к удалённому WPS сервису указанная пользователем таблица конвертируется с помощью утилиты `pgsql2shp`, входящей в стандартный пакет СУБД PostgreSQL, в файлы формата SHP. Далее сервису передаются HTTP ссылки на сконвертированные данные в соответствии со стандартом WPS.

Шаблон содержит подключение к сетевому диску, содержащему базовые пространственные данные. Через сетевой диск производится передача данных между сервисами, находящимися в облачной среде, что позволяет значительно ускорить передачу данных.

Разработка таблиц для хранения данных пользователей на основе структурных спецификаций. Геопортал обеспечивает создание, ввод и редактирование разных по структуре пользовательских таблиц. Структура таблиц (список полей, их названия, типы и т.д.) описывается структурными спецификациями. На основе структурных спецификаций создаются таблицы БД, генерируется пользовательский интерфейс и определяется логика работы геопортала. Кроме того, структурные спецификации позволяют автоматически связывать таблицы и сервисы обработки данных, упростить обмен данными между геопорталами. При создании таблицы пользователь может выбрать заранее сформированные структурные спецификации и при необходимости дополнить.

Структурные спецификации таблицы в формате JSON содержат: название таблицы и описание набора атрибутов. Каждый атрибут в свою очередь специфицируется: названием, именем в базе данных, типом данных, единицами измерения (для числовых данных), элементом управления и его свойствами. Элементы управления реализуют специфичные методы и необходимы для формирования пользовательского интерфейса добавления, редактирования и отображения данных атрибута. Свойства элемента управления позволяют настраивать пользовательский интерфейс в зависимости от характеристик данных, например, единицы измерения для числовых данных или определять тип пространственных данных.

Структурные спецификации и метаданные (авторы, дата обновления, и т.д.) хранятся в каталоге таблиц. Структурные спецификации таблиц в рамках каталога упорядочиваются в виде иерархий и применяют механизмы наследования и полиморфизма в терминах объектно-ориентированного подхода. Для создания таблицы пользователю необходимо создать модель в виде структурных спецификаций с помощью редактора модели, либо использовать готовую модель входных данных некоторого сервиса.

Модификация пользовательского интерфейса. Окончательный этап создания геопортала заключается в изменении HTML разметки. Для этого применяются различные шаблонизаторы.

Заключение. Создание геопортала на основе типового в облачной среде производится значительно быстрее. После развертывания шаблона виртуальной машины запускается работающий геопортал с готовым набором функций. Специализация геопортала заключается в создании таблиц, модификации пользовательского интерфейса, определении доступа к данным и т.д. Создание геопорталов на основе предложенного подхода позволяет:

- 1) создавать реляционные таблицы с произвольным количеством атрибутов, в том числе в соответствии со структурными спецификациями;
- 2) организовать многопользовательский ввод данных с возможностью определения доступа и модерации;
- 3) использовать при вводе данных унифицированные классификаторы, в том числе базовые пространственные данные;
- 4) сформировать общий каталог данных, позволяющий осуществлять поиск информации. Каталог хранит структурные спецификации таблиц, что позволяет производить программный поиск данных;

- 5) применять сервисы обработки данных, реализующих стандарт WPS. Геопортал предоставляет регламентированный доступ к данным в виде сервиса. Созданные таблицы можно применять в качестве входных данных в WPS-сервисах без специальной публикации;
 - 6) использовать общую систему идентификации и авторизации пользователей.
Разработанная технология в настоящее время применена для создания следующих ИС:
 - 1) Инвентаризация и анализ фиторазнообразия Байкальской природной территории, адрес <http://biodiv.isc.irk.ru>;
 - 2) Геопортал исследования распространения иксодовых клещей, адрес <http://tbd.icc.ru>;
 - 3) Атлас байкальской природной территории, адрес <http://atlas.isc.irk.ru>.
- Работа выполнена при финансовой поддержке РФФИ (гранты №№ 16-57-44034-монг_а, 16-07-00411-а).*

ЛИТЕРАТУРА

- [1] eBird [Электронный ресурс] / by the Cornell Lab of Ornithology and National Audubon Society. Электрон. текст. и граф. дан. URL: <http://ebird.org> (Дата обращения: 06.06.2017).
- [2] Плантариум [Электронный ресурс]: открытый атлас и онлайн-определитель растений России и сопредельных стран. Электрон. текст. и граф. дан. URL: <http://www.plantarium.ru> (Дата обращения: 06.06.2017).
- [3] Global Biodiversity Information Facility [Электронный ресурс]: Free and Open Access to Biodiversity Data / GBIF. Электрон. дан. URL: <http://data.gbif.org> (Дата обращения: 06.06.2017).
- [4] Catalogue of Life [Электронный ресурс]: online database of the world's known species of animals, plants, fungi and micro-organisms / by Catalogue of Life. Электрон. дан. URL: <http://www.catalogueoflife.org> (Дата обращения: 06.06.2017).
- [5] The Plant List [Электронный ресурс]: a working list of all known plant species. Электрон. текст. и граф. дан. URL: <http://www.theplantlist.org> (Дата обращения: 06.06.2017).
- [6] The International Plant Names Index = IPNI [Электронный ресурс]. Электрон. дан. URL: <http://www.ipni.org> (Дата обращения: 06.06.2017).
- [7] Информационная поисковая система по фауне и флоре заповедников России [Электронный ресурс] / разработ. и поддерж. группой «Биоинформатики и моделирования биологических процессов» ИПЭЭ РАН. Электрон. дан. URL: <http://www.sevin.ru/natreserves> (Дата обращения: 06.06.2017).
- [8] ООПТ России [Электронный ресурс]: Информационно-аналитическая система. Электрон. дан. URL: <http://oopt.aari.ru> (Дата обращения: 06.06.2017).
- [9] Информационно-поисковые системы и БД Зоологического института РАН [Электронный ресурс]. Электрон. дан. и прогр. URL: http://www.zin.ru/proj_r.htm (Дата обращения: 06.06.2017).
- [10] Cryptogamic Russian Information System [Электронный ресурс] : online information system Электрон. дан. URL: <http://www.kpabg.ru/cris> (Дата обращения: 06.06.2017).
- [11] Mapserver [Электронный ресурс] // Mapserver Consortium [сайт]. URL: <http://mapserver.org> (дата обращения: 04.06.2017).
- [12] MapCache [Электронный ресурс] // Mapserver Consortium [сайт]. URL: <http://mapserver.org/mapcache> (дата обращения: 04.06.2017).
- [13] PostgreSQL [Электронный ресурс] // The PostgreSQL Global Development Group [сайт]. URL: <http://www.postgresql.org> (дата обращения: 04.06.2017).
- [14] OpenGIS Web Processing Service (WPS) Implementation Specification, v1.0.0. Release date: June 08, 2007. [Электронный ресурс] // Open Geospatial Consortium [сайт]. URL: <http://www.opengeospatial.org/standards/wps> (дата обращения: 04.06.2017).