

РАЗРАБОТКА КОМПОНЕНТА ПОДДЕРЖКИ ВЗАИМОДЕЙСТВИЯ СЕТИ КАТАЛОГОВ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ НА БАЗЕ ПРОТОКОЛА Z39.50

Ильченко П.Н.

Научный руководитель — профессор Маглинец Ю.А.

Сибирский федеральный университет

До недавнего времени большинство архивов спутниковых данных создавалось в расчете либо на их локальное использование, либо на использование специальных выделенных высокоскоростных каналов связи для доступа к ним. С развитием современных технологий хранения и поиска данных появилась возможность создания электронных архивов нового поколения, позволяющих организовать оперативную работу широкого круга удаленных пользователей со спутниковыми данными и результатами их обработки.

В настоящее время объём спутниковых данных дистанционного зондирования Земли значительно возрастает, в связи с чем возникают задачи их систематизации, структуризации и создания средств каталогизации и поиска. Определяющую роль начинают играть не столько традиционные вопросы развития инфраструктуры мониторинга и создания информационных систем поддержки исследований природной среды, сколько задачи разработки новых информационных технологий, исследования применения методов и средств активных информационных технологий.

На сегодняшний день в мире существует огромное количество хранилищ данных дистанционного зондирования Земли, значительная часть которых предоставляет свободный доступ к своим хранилищам. Данные охватывают огромный географический и временной диапазоны снимков и метаданных.

Одной из основных систем управления пространственной информацией является «GeoNetwork», которая включает более 50 крупнейших хранилищ данных дистанционного зондирования Земли и связанных метаданных. Она предоставляет собственное программное обеспечение, которое базируется на протоколе Z39.50 и представляет собой отдельный сервер Apache Tomcat на каждом узле со своим веб-интерфейсом и репозиторием снимков. Но инструментов для использования в сторонних проектах разработчики не предоставляют.

В связи с этим возникла потребность создания компонента взаимодействия отдельно взятого хранилища с остальными узлами сети по протоколу Z39.50.

Протокол Z39.50 определяет порядок взаимодействия клиента и сервера, процедуры поиска и извлечения информации из баз данных и форматы представления этой информации, но не определяет форматы хранения данных в конкретных БД, способы их индексации и процедуры функционирования различных СУБД, также он не определяет интерфейсы взаимодействия пользователя и клиента.

В идеологии Z39.50 в рамках одной схемы все базы данных совершенно одинаковы, несмотря на их физические различия по используемой СУБД, полям и синтаксису запросов. В Z39.50 для клиента не существует возможности определить, под управлением какой СУБД хранятся извлекаемые им данные. Можно отметить, что подобная информация клиенту не нужна, так как он работает всегда с одной и той же системой запросов и получает данные в одних и тех же форматах.

В 1995 году протокол Z39.50 был принят как стандарт ANSI, т.е. как американский национальный стандарт. Полное обозначение стандарта – ANSI/NISO Z39.50-1995.

В ноябре 1998 года протокол Z39.50 был принят как международный стандарт ISO-23950.

Протокол Z39.50 описывает сетевое взаимодействие субъектов в архитектуре «клиент-сервер». Однако это взаимодействие несколько отличается от классической архитектуры «клиент-сервер», в которой инициатором любого запроса может быть только клиент, а серверу всегда отводится пассивная роль ожидающего и отвечающего. Как будет видно ниже, в Z39.50 это не всегда так. Может быть поэтому разработчики протокола изменили терминологию, заменив термины «клиент» и «сервер» на термины «origin» и «target» соответственно. За редким исключением понятия «клиент»-«origin» и «сервер»-«target» совпадают.

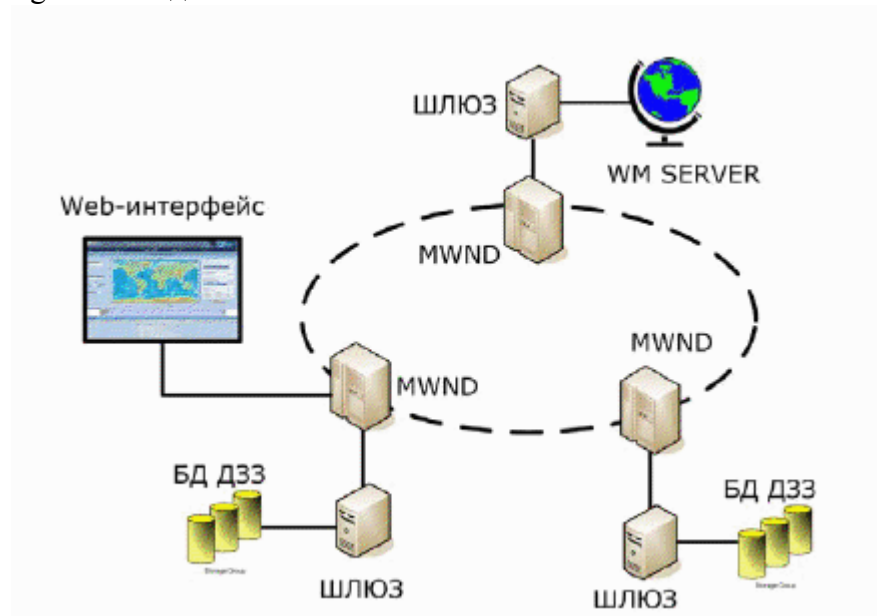


Рис. 1.

Основная функция такой системы – создание единой информационной среды, предоставление единой точки доступа для поиска и получения информации по всем архивам и каталогам, зарегистрированным в системе, независимо от их географического расположения и внутреннего формата данных. Распределенная информационная система (рис. 1) представляет собой сеть узлов, так называемых Middleware Nodes (MWNDs). Каждый из узлов содержит метаданные коллекций, описывающие данные, доступ к которым предоставляется через Шлюзы (Gateways). Пользователь для поиска информации обращается к системе через WWW-сервер на одном из узлов MWND, где формируется распределенный запрос, рассылаемый релевантным шлюзам. Результаты выполнения запроса объединяются и возвращаются пользователю.

Шлюз представляет собой промежуточное звено между узлом MWND и каталогом или архивом поставщика данных. Шлюз обменивается сообщениями по протоколу CIP (Catalogue Interoperability Protocol) с узлом MWND, получая от него поисковые запросы и возвращая ему результаты поиска.

Запросы передаются транслятору RDBMS, который преобразует их в запросы для непосредственного поиска в базе данных, либо формирует запрос для обращения к предоставляемому сервису (Web Map Server и т.п.).

В основе распределенной информационной системы лежит протокол CIP, регламентирующий правила взаимодействия пользователей и каталогов данных ДЗЗ. Для поддержки одновременного доступа пользователя ко множеству каталогов используется трехуровневый принцип распределения запросов. Пользователь через web-интерфейс задает поисковый запрос и посылает его к узлу MWND, который, в свою очередь, пере-

направляет запрос множеству серверов каталогов данных. Сервера, имеющие данные, удовлетворяющие критериям поиска, возвращают ответы на узел MWND, через который пользователь получает объединенный результат поиска.

Компонент должен представлять собой шлюз между АСПКС и внешними хранилищами спутниковых данных. Взаимодействуя с внешними репозиториями по протоколу Z39.50, он должен возвращать данные локальному сервису в определенном формате, попутно взаимодействуя с локальной базой данных по уже существующим интерфейсам.

Разработка компонента взаимодействия системы АСПКС с другими внешними репозиториями спутниковых данных на базе протокола Z39.50 позволит реализовать доступ к данным дистанционного зондирования Земли большой номенклатуры космических аппаратов, а также с более широким территориальным и временным охватом.