

Применение инструментария IARnet для интеграции аналитических ГИС и распределенных информационных ресурсов*

О. В. Сухорослов

*Институт системного анализа Российской академии наук
(ИСА РАН)*

Одной из главных целей проекта «Электронная Земля» [1] является создание единого информационного поля в области наук о Земле, а также соответствующей сетевой и программной инфраструктуры, позволяющей интегрировать разнородные информационные ресурсы участников проекта. В статье описываются проведенные в рамках данного проекта работы по интеграции информационно-аналитических систем КОМПАС и ГеоПроцессор с распределенными информационными ресурсами с применением инструментария IARnet [2].

1. Интеграция системы КОМПАС с распределенными информационными ресурсами

1.1. Система КОМПАС

Информационно-аналитическая система КОМПАС-III (COMPASS: Cartography Online Modeling, Presentation and Analysis System) [3], разрабатываемая в ИППИ РАН, предназначена для интерактивного представления, анализа и моделирования тематической информации, которая может относиться к любым географическим объектам. Система представляет информацию в виде карт (картограмм), диаграмм и таблиц. Данные, с которыми работает пользователь системы оформлены в виде

* Работа выполнена при поддержке фонда РФФИ (грант № 05-07-90182-в), Президиума РАН (программа фундаментальных исследований 15П, проекты 1.1 и 1.4) и Федерального агентства по науке и инновациям (государственный контракт № 02.514.11.4053).

набора файлов (так называемый ГИС-проект), содержащих картографические и атрибутивные данные, связи между данными и параметры представления информации на экране.

Система КОМПАС реализована на языке Java и включена в Web-страницу в виде апплета. При загрузке такой страницы на компьютер пользователя автоматически загружается сама система, а также определенный ГИС-проект, связанный с данной страницей. После этого обращений к серверу не происходит, и используются только ресурсы компьютера пользователя. Именно это обстоятельство позволяет достичь высокой интерактивности при визуализации и анализе информации.

1.2. Требования к интеграции системы КОМПАС с распределенными информационными ресурсами

Существовавшая на момент начала работ (2005 г.) третья версия системы КОМПАС требовала, чтобы все данные, с которыми работает пользователь, находились на том же сервере, что и апплет системы (см. рис. 1). Это обстоятельство делало невозможным работу с данными, размещенными их владельцами на других машинах в сети. Фактически, для того, чтобы пользователи системы могли работать с подобными данными, требовалось вручную перенести их на сервер, где работает КОМПАС, и подготовить соответствующую Web-страницу. Эти мероприятия требовали непосредственного участия владельцев данных и разработчиков системы, при этом после переноса данных их владельцы теряли над ними контроль.

С точки зрения владельцев данных, наиболее приемлемым вариантом является размещение данных на машине, доступом к которой

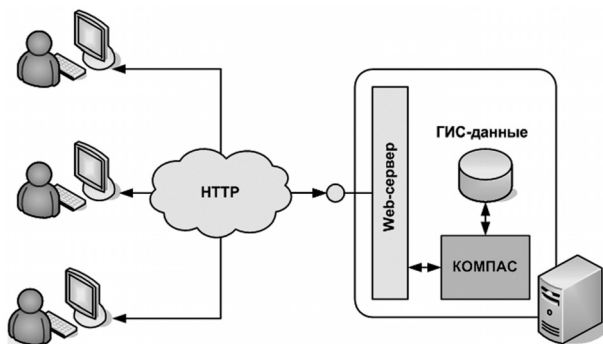


Рис. 1. Изначальная схема работы системы КОМПАС

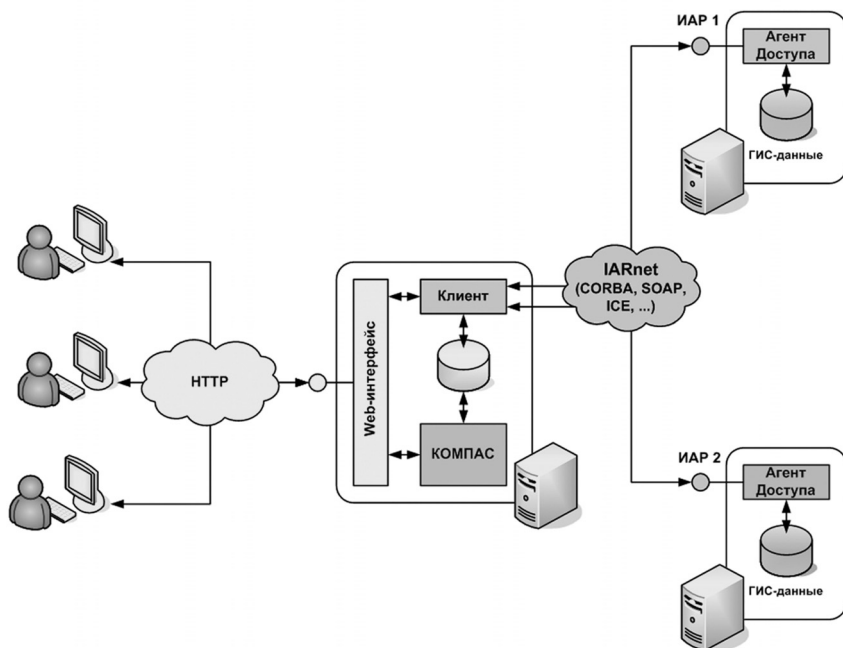


Рис. 2. Схема интеграции распределенных данных и ГИС КОМПАС на основе системы IARnet

управляют они сами, — например, на сервере соответствующего института РАН. При этом владельцы данных не теряют контроль над данными и могут оперативно их обновлять. Подобный вариант, без использования каких-либо дополнительных средств, был реализуем только путем установки системы КОМПАС на все машины, на которых размещены данные, что в свою очередь противоречит интересам разработчиков системы, которые не хотят терять контроль над своим программным продуктом.

Таким образом, требовалась разработка и использование специальных средств для интеграции удаленных данных с системой КОМПАС. В соответствии с указанными обстоятельствами были сформулированы следующие требования:

- Пользователи системы КОМПАС должны иметь возможность работать посредством системы с данными (информационными ресурсами), размещенными в любом месте сети.

- Владельцы данных должны иметь возможность разместить данные на любой доступной им машине в сети и полностью управлять доступом к данным.
- Все операции, связанные с интеграцией системы КОМПАС и удаленных данных, должны быть автоматизированы и не должны требовать ручного вмешательства пользователей. В идеале, пользователи должны работать с удаленными данными (размещенными вне сервера системы КОМПАС) в точности так же, как с локальными данными (размещенными на сервере системы КОМПАС).

1.3. Реализация интеграции системы КОМПАС с распределенными информационными ресурсами

В соответствии с заданными требованиями был реализован пилотный вариант программных средств для интеграции распределенных данных и ГИС КОМПАС на базе инструментария IARnet (см. рис. 2). Данные средства включают в себя:

- Агент доступа к ГИС-проектам системы КОМПАС.
- Web-интерфейс, предназначенный для организации работы пользователя с распределенными ГИС-проектами в системе КОМПАС.

1.3.1. Агент доступа к ГИС-проектам системы КОМПАС

В соответствии с архитектурой IARnet, наборы данных для КОМПАС были оформлены в виде информационно-алгоритмических ресурсов, снабженных специальными агентами доступа. Каждый агент предоставляет доступ к определенному набору ГИС-проектов, каждый из которых представляет собой набор файлов.

По запросу клиента агент может:

- передать клиенту список всех имеющихся у него проектов с их идентификаторами и описаниями;
- передать клиенту ГИС-проект (упакованный в zip-архив) по его идентификатору.

В соответствии с этим, интерфейс агента доступа содержит две операции:

- [http://dcs.isa.ru/iarnet/CompassDataResource#listData\(\):string\[\]](http://dcs.isa.ru/iarnet/CompassDataResource#listData():string[]);
- [http://dcs.isa.ru/iarnet/CompassDataResource#getData\(string\):byte\[\]](http://dcs.isa.ru/iarnet/CompassDataResource#getData(string):byte[]).

При инициализации агенту доступа передается путь к директории с обслуживаемыми ГИС-проектами, где каждый проект размещен внутри отдельной директории с именем «\$Идентификатор_проекта».

Агент доступа реализован на языке Java и развертывается в контейнере ИАР IceContainer на основе технологии Ice [4].

1.3.2. Интегрированные информационные ресурсы

В настоящее время в рамках созданной системы организован доступ к следующим ГИС-проектам:

- Экономические показатели по России (Министерство экономики и развития торговли).
- Экологическая ситуация по округам Москвы.
- Природные ресурсы и окружающая среда России.
- Интегральная карта природных опасностей и рисков Алтая.
- Интегральная карта природных опасностей и рисков Азии.
- Интегральная карта природных опасностей и рисков России.
- Социально-экономическая информация по субъектам РФ.

1.3.3. Web-интерфейс пользователя

Для работы пользователей с удаленными информационными ресурсами на сервере с системой КОМПАС был создан специальный Web-интерфейс (рис. 3), который позволяет пользователю:

- просматривать список всех доступных ГИС-проектов;
- открывать выбранный проект в системе КОМПАС в отдельном окне браузера.

Указанный Web-интерфейс реализован на языке Java Server Pages (JSP) и использует для доступа к ресурсам клиентское приложение, написанное на языке Java. При инициализации Web-приложения происходит считывание из конфигурационного файла списка ссылок на доступные ресурсы с данными, которые используются далее для вызова ресурсов через клиентское приложение.

Рассмотрим сценарий отображения списка доступных в сети ГИС-проектов:

- Пользователь заходит с помощью браузера на главную страницу Web-интерфейса.
- Web-приложение через клиентское приложение последовательно запрашивает список проектов с каждого известного ей ресурса (при этом производятся удаленные вызовы агентов доступа).

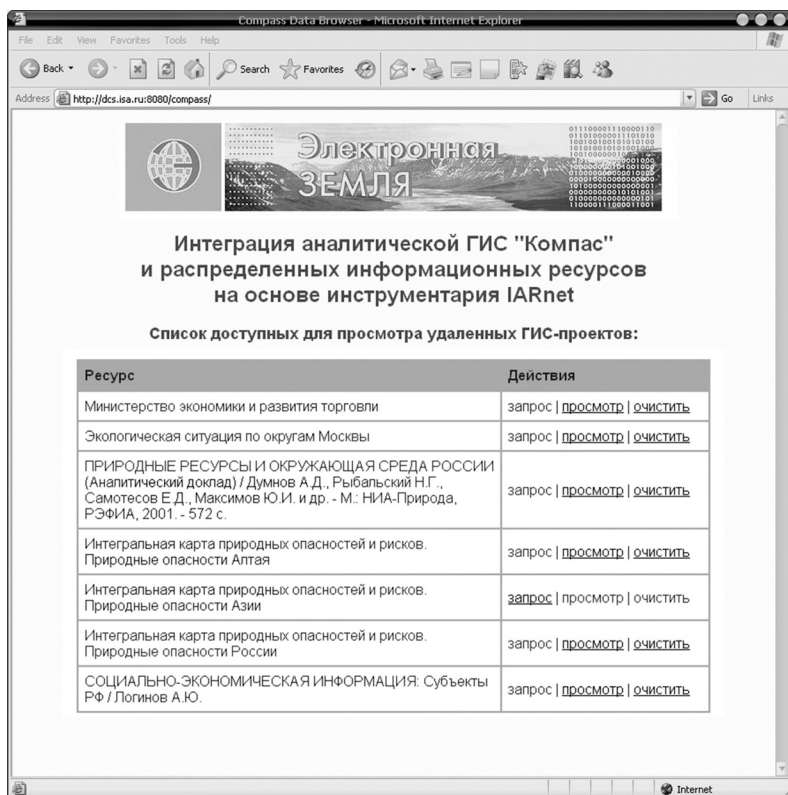


Рис. 3. Web-интерфейс пользователя

- Web-приложение формирует сводный список всех доступных ГИС-проектов с их названиями и описаниями и выдает его браузеру пользователя.

Рассмотрим подробно сценарий работы пользователя в ГИС-проектом через Web-интерфейс:

- Пользователь выбирает в списке проектов интересующие его данные и нажимает ссылку «запрос» справа от описания проекта.
- По запросу пользователя инициализируется передача данных ГИС-проекта на сервер:
 - web-приложение определяет ресурс, на котором находятся требуемые данные, и запрашивает их у ресурса через клиентское приложение;

- ресурс передает данные проекта клиентскому приложению, которое распаковывает их в выделенной директории на сервере и автоматически создает Web-страницу для просмотра данных в системе КОМПАС;
- при успешной загрузке данных Web-приложение обновляет список ГИС-проектов — активирует (делает доступной) ссылку «просмотр» и деактивирует ссылку «запрос» у рассматриваемого проекта.
- Пользователь нажимает на ссылку «просмотр» у рассматриваемого проекта, после чего в новом окне открывается сгенерированная ранее Web-страница для просмотра данных в системе КОМПАС.
- После окончания работы с данными пользователь закрывает окно системы КОМПАС.

В текущей реализации полученные от агента данные «кэшируются» на сервере бесконечно. В то же время, владельцы данных могут изменить содержание оригинальных данных в любой момент. Для того чтобы обновить закэшированные данные, используется следующий сценарий:

- Пользователь нажимает ссылку «удалить» (активирована у всех загруженных проектов) у требуемого проекта, после чего происходит удаление локально закэшированной копии проекта.
- Пользователь нажимает ссылку «запрос», после чего происходит повторная загрузка данных проекта с соответствующего ресурса.

2. Интеграция системы ГеоПроцессор с распределенными информационными ресурсами

2.1. Система ГеоПроцессор

Информационно-аналитическая система ГеоПроцессор (GeoProcessor) [5], разрабатываемая в ИППИ РАН, предназначена для интерактивного представления и анализа пространственно-временной геоинформации. Аналогично системе КОМПАС, данные, с которыми работает пользователь ГеоПроцессора оформлены в виде набора файлов (так называемый ГИС-проект), содержащих картографические и атрибутивные данные, связи между данными и параметры представления информа-

ции на экране. Система реализована на языке Java и включена в Web-страницу в виде апплета. При загрузке такой страницы на компьютер пользователя автоматически загружается сама система, а также определенный ГИС-проект, связанный с данной страницей.

В отличие от системы КОМПАС, при разработке GeoПроцессора сразу была заложена возможность загрузки удаленных ГИС-проектов, то есть проектов, не находящихся на одной машине с апплетом GeoПроцессора. В текущей версии системы поддерживается загрузка удаленных ГИС-проектов по протоколу HTTP. Для этого требуется заметить в Web-странице с апплетом локальный путь к главному файлу проекта (ROOT.INI) на URL этого файла на удаленном HTTP-сервере. Обнаружив вместо локального пути URL, система автоматически загружает файлы проекта с указанного сервера. Отметим, что для работы описанного механизма, необходима настройка политики безопасности Java, не допускающей по умолчанию соединение апплета с удаленными серверами, за исключением того сервера, с которого был загружен апплет.

2.2. Реализация интеграции системы GeoПроцессор с распределенными информационными ресурсами

Поскольку система GeoПроцессор уже поддерживает загрузку удаленных ГИС-проектов через протокол HTTP, было решено расширить эти возможности путем встраивания в систему поддержки удаленного доступа к ГИС-проектам через IARnet. Для этого были реализованы следующие программные компоненты:

- Агент доступа к ГИС-проектам GeoПроцессора.
- Библиотека IARnetGeoLib, подключаемая к GeoПроцессору и реализующая доступ к удаленным ГИС-проектам через указанных выше агентов доступа.

2.2.1. Агент доступа к ГИС-проектам GeoПроцессора

Поскольку при открытии ГИС-проекта GeoПроцессор проводит считывание каждого файла проекта по отдельности, при реализации агента доступа к ГИС-проекту необходимо обеспечить удаленный доступ к данным на уровне отдельных файлов проекта.

В текущей реализации, интерфейс агента доступа к ГИС-проектам GeoПроцессора (далее, просто агент доступа) содержит единственную операцию:


```
http://dcs.isa.ru/iarnet/GeoDataResource  
#getGeoData(string):byte[].
```

При вызове операции, в качестве аргумента клиент передает строковый идентификатор файла ГИС-проекта. Идентификатор файла имеет вид: «\$имя_проекта/\$имя_файла». В случае, если файл с указанным идентификатором существует, агент возвращает клиенту содержимое файла в виде байтового потока.

При инициализации агенту доступа передается путь к директории с обслуживаемыми ГИС-проектами, где каждый проект размещен внутри отдельной директории с именем «\$имя_проекта».

Агент доступа был реализован на языке Java и развертывается в контейнере ИАР IceContainer.

2.2.2. Библиотека IARnetGeoLib

Подключаемая к ГеоПроцессору библиотека IARnetGeoLib позволяет загружать удаленные ГИС-проекты, доступ к которым реализован при помощи описанного выше агента доступа. Библиотека реализована на языке Java и содержит единственную операцию:

```
InputStream getStreamData(String address).
```

В качестве аргумента данной операции передается адрес файла ГИС-проекта, который имеет вид: «iarnetdata:\$agent_reference:\$file_id», где \$agent_reference — IARnet-ссылка на агент доступа, а \$file_id — описанный ранее идентификатор файла.

При вызове операции библиотека производит удаленный вызов соответствующего агента доступа, передавая ему указанный идентификатор файла, получает в ответ байтовый поток с содержимым файла и возвращает его приложению в виде объекта InputStream.

2.2.3. Модификация стартовой Web-страницы ГеоПроцессора

Для открытия в ГеоПроцессоре удаленного ГИС-проекта, доступ к которому реализован при помощи описанного выше агента доступа, в стартовой Web-странице необходимо в качестве значения параметра data указать адрес главного файла проекта (ROOT.INI) в описанном выше формате:

```
<param name="data" value="$root_ini_address">.
```

Отметим, что модификация самих файлов проекта, включая ROOT.INI, при переходе с локального к удаленному режиму доступа не требуется.

2.3. Дальнейшее развитие системы

Текущая реализации механизма загрузки удаленного ГИС-проекта не является идеальной. При загрузке проекта GeoПроцессор проводит считывание каждого файла проекта по отдельности, что приводит к серии удаленных вызовов агента доступа, каждый из которых связан с передачей только одного файла. Данный подход обязателен для загрузки файлов по протоколу HTTP или в экзотических случаях, когда файлы проекта хранятся на нескольких серверах. В случае IARnet, накладные расходы по загрузке проекта можно существенно снизить, получив все содержимое проекта за один удаленный вызов, так же как это сделано для системы КОМПАС. При этом данные передаются в сжатом виде (zip-компрессия), снижая также потребляемый сетевой трафик. Совместимость со стандартным механизмом загрузки проекта может быть достигнута путем временного кэширования файлов проекта после их загрузки библиотекой IARnetGeoLib с последующей выдачей их содержимого из кэша.

Другим возможным направлением развития системы является обеспечение защищенного доступа к данным. Агент может контролировать доступ к ГИС-проектам в соответствии с описанными владельцем ресурса правами доступа, например — разрешать чтение файлов проекта только определенным пользователям. Для идентификации пользователей могут применяться их учетные записи на центральном портале проекта «Электронная Земля» или специально созданные цифровые сертификаты формата X.509. В обоих случаях необходима соответствующая модификация GeoПроцессора для поддержки ввода реквизитов пользователя и его аутентификации с тем, чтобы система могла обращаться к ресурсам от имени пользователя и проходить надлежащую процедуру авторизации.

Литература

1. Центральный портал проекта «Электронная Земля» Президиума РАН / <http://earth.viniti.ru>
2. Емельянов С. В., Афанасьев А. П., Волошинов В. В., Гринберг Я. Р., Кривцов В. Е., Сухорослов О. В. Реализация Grid-вычислений в среде IARnet // Информационные технологии и вычислительные системы. М.: Институт микропроцессорных вычислительных систем РАН, 2005. № 2.
3. Система КОМПАС / <http://gis.iitp.ru/compass3>
4. Сухорослов О. В. Промежуточное программное обеспечение Ice (в наст. сборнике).
5. Система GeoProcessor / <http://gis.iitp.ru/geoproc>