

УДК 004.9:502

С. В. ПАВЛОВ, Р. З. ХАМИТОВ, А. Б. НИКИТИН

СТРУКТУРА РАЗНОРОДНОЙ ТЕРРИТОРИАЛЬНО-РАСПРЕДЕЛЕННОЙ ПРОСТРАНСТВЕННОЙ ИНФОРМАЦИИ ПРИ СОЗДАНИИ ЕДИНОЙ ГЕОИНФОРМАЦИОННОЙ СИСТЕМЫ РОСВОДРЕСУРСОВ

Рассматриваются вопросы проектирования распределенной геоинформационной системы Федерального агентства водных ресурсов. Обоснована необходимость использования геоинформационных систем как инструмента для получения оперативной, точной и целостной информации о водных и связанных с ними объектах. Обосновывается необходимость и приводится один из возможных подходов к формальному описанию структуры распределенных пространственных данных, рассматриваются вопросы их практической реализации. *Геоинформационная система; распределенная обработка; пространственная информация; многомерные информационные объекты*

ВВЕДЕНИЕ

На сегодняшний день за сбор, хранение, обработку и распространение данных, характеризующих объекты водной отрасли, отвечают многие государственные организации, такие как Министерство природных ресурсов, Федеральное агентство водных ресурсов, Росгидромет и др.

При этом существует разница в программном обеспечении, используемом для управления данными, а содержание баз данных по качественным характеристикам объектов водной отрасли в целом неудовлетворительно [1]. Имеющиеся различия в программном обеспечении, в структуре и содержании баз данных не позволяют осуществить комплексное применение всего спектра современных технологий построения крупных корпоративных информационных систем для обеспечения государства и общества полной и достоверной пространственной и атрибутивной информацией (справочной и аналитической) о водных объектах Российской Федерации (РФ).

Основная цель деятельности Федерального агентства водных ресурсов (Росводресурсов) заключается в управлении водными ресурсами страны. Для достижения этой цели необходима полная, своевременная, непротиворечивая информация о текущем и прогнозируемом состоянии водных ресурсов. Также для успешного решения задач управления

необходима информация об иных объектах и природных условиях по территориям, на которых расположены водные объекты, о пользователях, а также о показателях, характеризующих основные аспекты (финансовые, организационные, технические и экологические) осуществляемой управленческой деятельности.

Особенностью рассматриваемых водных, водохозяйственных и других объектов (реки, озера, каналы, водохранилища, гидротехнические сооружения) является их существенная протяженность и распределенность по всей территории России и сопредельных государств. Причем наличие информации о точном местоположении объектов, их взаимном расположении и взаимосвязи существенно влияет на качество управленческих решений, принимаемых и реализуемых на различных уровнях.

Так как управление в структуре Росводресурсов реализуется на различных уровнях: федеральном, бассейновом, территориальном, местном, то и соответствующее информационное обеспечение должно быть построено по иерархическому принципу с разной степенью детализации и обобщения информации на каждом уровне.

В связи с этим при создании современного информационного обеспечения предприятий, учреждений и органов Росводресурсов всех уровней и особенно при их объедине-

нии (интеграции) в единую корпоративную информационную систему в качестве технологической основы выбраны геоинформационные системы.

1. ПРИНЦИПЫ ИНТЕГРАЦИИ РАЗНОРОДНОЙ И ТЕРРИТОРИАЛЬНО-РАСПРЕДЕЛЕННОЙ ПРОСТРАНСТВЕННОЙ ИНФОРМАЦИИ

Процесс управления водными ресурсами имеет 3 уровня управления (рис. 1) [1].



Рис. 1. Структура взаимодействия частей распределенной ГИС Росводресурсов по уровням управления

Создание геоинформационной системы (ГИС) Росводресурсов должно вестись применительно к этим трем уровням управления: территориальному, бассейновому и центральному. При этом в зависимости от уровня управления меняется как состав решаемых задач, так и состав необходимой пространственной и атрибутивной информации. Помимо логической иерархии по решаемым на разных уровнях управления задачам и детализации информации, необходима также территориальная распределенность компонент геоинформационной системы ввиду территориальной распределенности подразделений Росводресурсов.

На рис. 1 представлена схема взаимодействия частей распределенной геоинформационной системы Росводресурсов по уровням управления.

Здесь GIS_{Φ} — геоинформационная система, содержащая в себе информацию по всей территории Российской Федерации; $\{GIS_i\}$ — множество геоинформационных систем уровня бассейнового водного управления (БВУ), где $i = \overline{1, N}$. N — количество БВУ; $\{GIS_j\}$ — множество геоинформационных систем территориального уровня для

каждого БВУ, где $i = \overline{1, N}$ и $j = \overline{1, M_i}$, N — количество БВУ, M_i — количество территориальных отделов для каждого БВУ.

Говоря о создании распределенной геоинформационной системы, подразумевается распределение частей системы как по их хранению в различных базах данных, так и по распределению хранения информации о различных частях территории РФ.

При проектировании распределенных информационных систем существует несколько стратегий. Целью оптимального проектирования является синтез распределенной системы, обеспечивающей экстремум критерия эффективности, с учетом ограничения на технико-экономические характеристики разрабатываемой системы. Например, стоимость и время разработки, отладки, тестирования и эксплуатации системы; затраты машинной памяти и машинного времени для выполнения задач; избыточность; надежность, время наработки на отказ; затраты на верифицирование результатов работы модулей системы и т. д.

В данном случае мы рассматриваем распределенные системы, имеющие иерархическую структуру (имеющие центр), в соответствии со структурой управления водными ресурсами. В соответствии с рис. 1 можно построить следующую иерархию распределенной ГИС Росводресурсов (рис. 2) (без учета территориального уровня).

На данном этапе создания распределенной геоинформационной системы Росводресурсов предлагается остановиться на бассейновом уровне при проектировании и разработке. При этом на рис. 2 показаны базовые масштабы для ГИС каждого уровня. Следует отметить, что вопрос с базовым масштабом для бассейнового уровня не решается однозначно. Например, в Восточной Сибири с малой плотностью населения и хозяйственных построек и большой площадью территории имеет смысл использовать масштаб 1:500 000. В то же время в Европейской части страны для некоторых БВУ возможно использовать картографическую основу масштаба 1:100 000.

В достаточно общем виде задача создания распределенной геоинформационной системы может быть сформулирована следующим образом: необходимо построить отображения F_i , переводящие геоинформационные системы GIS_i в GIS_{Φ} :

$$F_i : GIS_i \rightarrow GIS_{\Phi}.$$

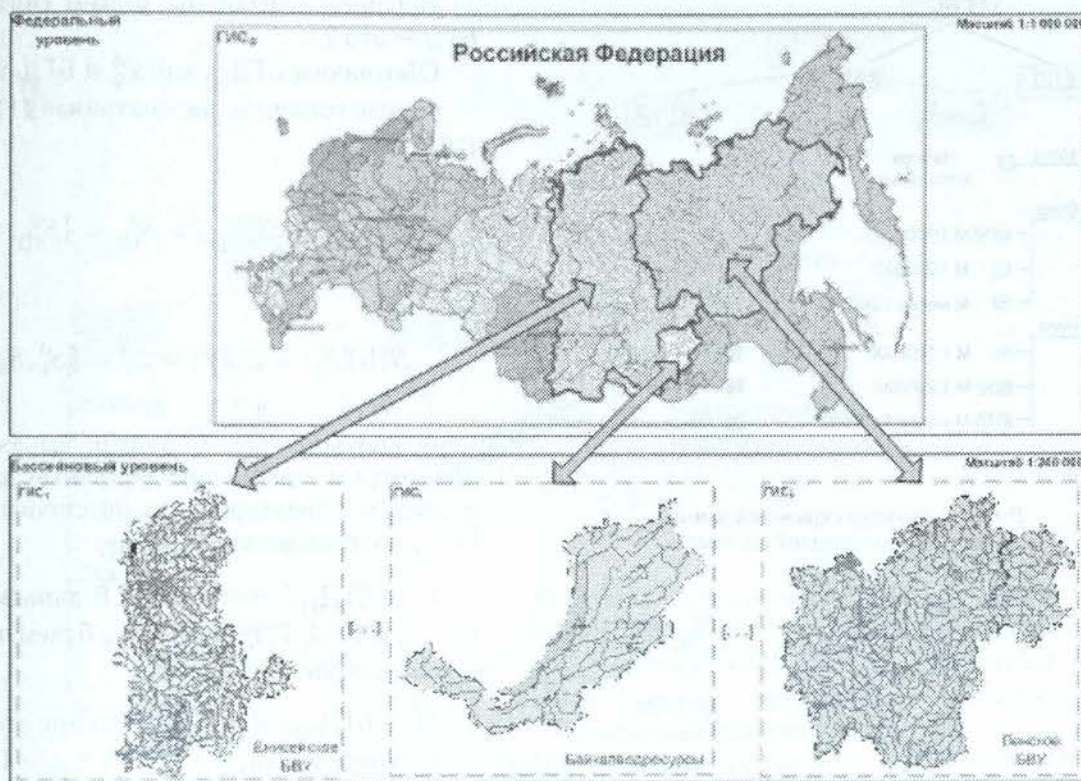


Рис. 2

Геоинформационную систему (ГИС) можно определить, как $ГИС = \{БГД, БАД, P\}$, где БГД – база географических данных, БАД – база атрибутивных данных, P – множество процедур обработки данных.

Соответственно, $ГИС_{\Phi} = \{БГД_{\Phi}, БАД_{\Phi}, P\}$, где $БГД_{\Phi}$ – база географических данных федерального уровня, $БАД_{\Phi}$ – база атрибутивных данных федерального уровня, P – множество процедур обработки данных.

При этом каждый элемент множества $\{ГИС_i\}$ (геоинформационная система бассейнового уровня) можно обозначить как:

$$ГИС_i = \{БГД_i, БАД_i, P\},$$

где $БГД_i$ – база географических данных, $БАД_i$ – база атрибутивных данных, P – множество процедур обработки данных.

В данном случае мы считаем, что множество процедур обработки пространственных и атрибутивных данных на каждом уровне ГИС Росводресурсов совпадает (примеры: распространение загрязнений, построение зон затоплений).

Характерной особенностью географических данных, используемых в ГИС, является их высокая структурированность. В отличие от обычной реляционной БД, не накладывающей никаких ограничений на структуру хранимых в ней отношений, все простран-

ственные объекты (слои), хранящиеся в БГД, имеют минимальный общий набор атрибутов, характеризующий их положение в пространстве и другие геометрические характеристики (например, длину, или площадь). Слои БГД подразделяются по степени детализации их геометрических характеристик, определяемой масштабом соответствующей карты. При этом одни и те же географические объекты могут быть представлены слоями разных типов (точечными, линейными и полигональными) в зависимости от степени детализации пространственных данных.

Соответственно, структуру хранения данных в геоинформационной системе можно представить следующим образом (рис. 3). При этом в БАД геоинформационной системы атрибутивные данные хранятся обычным для реляционных СУБД образом.

Для формального описания структуры пространственных данных целесообразно использовать концепцию многомерных информационных объектов (МИО). При этом существует механизм совместного описания пространственных данных, представленных МИО различной размерности, и атрибутивных данных, представляющих из себя реляционные таблицы и связывающие их отношения [2].

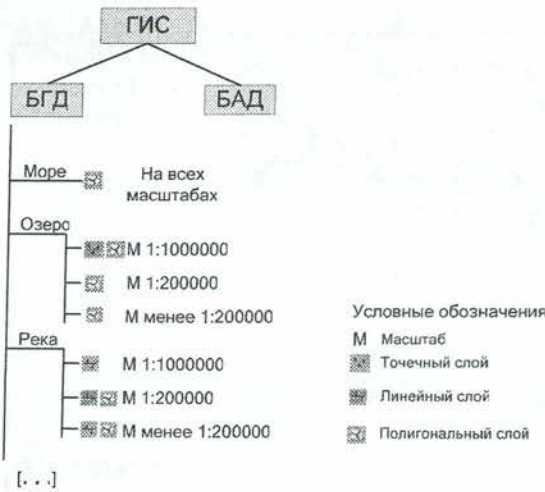


Рис. 3. Структура хранения данных в геоинформационной системе

Множество $\{ГИС_i\}$ неоднородно, потому что элементы $ГИС_i$ могут отличаться друг от друга и от $ГИС_\Phi$ структурами хранения пространственной информации. Поэтому будут отличаться друг от друга и отображения F_i .

Вспользуемся аппаратом МИО из [2] для формального описания видов элементов $ГИС_i$. Используем понятие схемы МИО для описания внутренней структуры МИО и соответственно БГД каждого уровня ГИС Ресурсов.

Схема МИО $S(T^n) = S^n$ представляет собой множество, значения которого задают порядок вхождения МИО размерности $n-1$ (с учетом порядка вхождения в них МИО меньших размерностей) в МИО размерности n . В общем случае схема МИО размерности n будет определяться соотношением

$$S^n = S(T^n) = \{S(T^{n-1}), S_n\} = \{S^{n-1}, S_n\} = \dots = \{S_1, S_2, \dots, S_{n-1}, S_n\} \quad (1)$$

где

$$S^{n-1} = \{S_1, S_2, \dots, S_{n-1}\},$$

$$S^{n-2} = \{S_1, S_2, \dots, S_{n-2}\},$$

...

$$S^2 = \{S_1, S_2\}$$

являются схемами МИО размерности $(n-1)$, $(n-2)$, ..., 2, входящих в T^n , а $S_1, S_2, \dots, S_{n-1}, S_n$ — элементами схемы МИО размерности n ; T^n — множество МИО размерности $(n-1)$.

При этом, если

$$S(T_i^n) = S(T_j^n),$$

то объекты T_i^n, T_j^n являются однотипными. В противном случае они являются разнотипными.

В нашем случае мы можем считать, что БГД — это T^4 .

Обозначим БГД $_\Phi$ как T_Φ^4 и БГД $_i$ как T_i^4 .

Соответственно, рассматривая (1), можно записать

$$S(\text{БГД}_\Phi) = S(T_\Phi^4) = S_\Phi^4 = \{S_\Phi^3, S_{\Phi 4}\};$$

$$S(\text{БГД}_i) = S(T_i^4) = S_i^4 = \{S_i^3, S_{i4}\}.$$

В соответствии с этими выражениями рассмотрим следующие основные варианты и введем характеристику различных типов $ГИС_i$ по отношению к $ГИС_\Phi$:

1) $S(\text{БГД}_\Phi) = S(\text{БГД}_i)$. В данном случае БГД $_\Phi$ и БГД $_i$, $ГИС_\Phi$ и $ГИС_i$ будем называть однотипными.

2) $S(\text{БГД}_\Phi) \neq S(\text{БГД}_i)$, но при этом $S_\Phi^3 = S_i^3$. Тогда БГД $_\Phi$ и БГД $_i$, $ГИС_\Phi$ и $ГИС_i$ будем называть разнотипными по масштабам.

3) $S(\text{БГД}_\Phi) \neq S(\text{БГД}_i)$ и $S_\Phi^3 \neq S_i^3$. Тогда БГД $_\Phi$ и БГД $_i$, $ГИС_\Phi$ и $ГИС_i$ будем называть разнотипными по составу.

Фрагменты структур БГД для $ГИС_i$ различных типов по отношению к $ГИС_\Phi$ приведены на рис. 4.

Построение отношений F_i должно вестись в соответствии с приведенной классификацией элементов $ГИС_i$. При этом очевидно, что случай однотипности элемента $ГИС_i$ по отношению к $ГИС_\Phi$ является наиболее простым, в то время как случай разнотипности по составу является наиболее сложным.

В соответствие с разделением множества $\{ГИС_i\}$ можно ввести разделение отображений F_i на следующие группы:

1) $F_i^{\text{одн}}$ в случае, когда $ГИС_i$ и $ГИС_\Phi$ являются однотипными.

2) $F_i^{\text{рм}}$ в случае, когда $ГИС_i$ и $ГИС_\Phi$ являются разнотипными по масштабам.

3) $F_i^{\text{рс}}$ в случае, когда $ГИС_i$ и $ГИС_\Phi$ являются разнотипными по составу.

Таким образом, задача построения отображений F_i должна сводиться к формальному описанию пространственных данных на каждом уровне системы и построению правил ее преобразования.

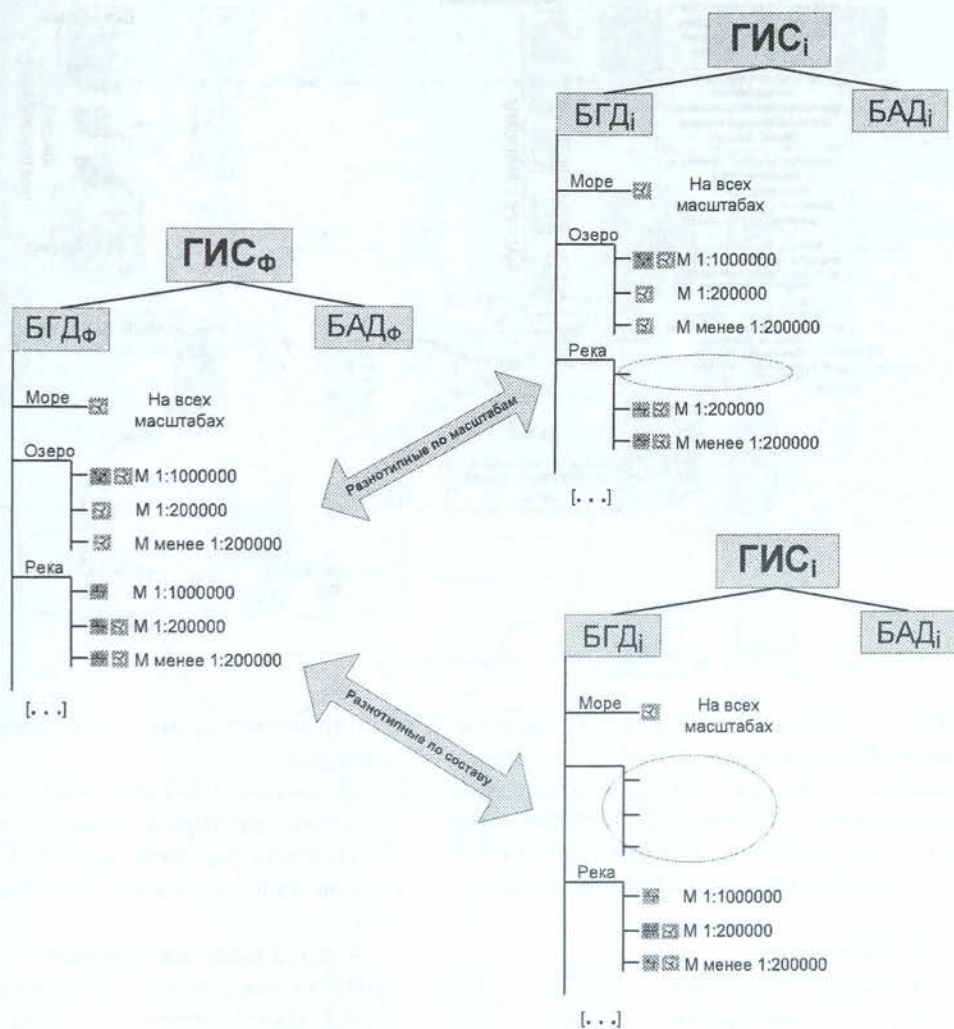


Рис. 4. Примеры разнотипных структур GIS_i по отношению к GIS_φ

До сих пор рассматривалась только пространственная составляющая (БГД) геоинформационной системы. В случае с атрибутивной составляющей геоинформационной системы возможно применение аппарата реляционной алгебры для описания структур атрибутивных данных для элементов распределенной ГИС Росводресурсов. Также возможно применение расширенных МИО [2] для совместного описания атрибутивных и пространственных данных.

При этом при проектировании ГИС Росводресурсов необходимо учитывать наличие на всех уровнях управления водными ресурсами систем обработки атрибутивной информации, которые необходимо интегрировать с ГИС. Интеграция может проходить на нескольких уровнях: на уровне клиентских приложений и на уровне серверных компонентов, обеспечивающих централизованное хранение географических данных.

2. ПРИМЕРЫ РЕАЛИЗАЦИИ

Типовая структура элемента ГИС Росводресурсов приведена на рис. 5.

ArcSDE является ГИС-шлюзом к реляционным базам данных и позволяет СУБД полностью интегрироваться с ГИС. ArcSDE является ключевой компонентой в многопользовательской ГИС, поскольку позволяет связать стандарты технологий реляционных баз данных со стандартами технологий ГИС.

Использование ArcSDE даст следующие преимущества:

- поддержка параллельного многопользовательского редактирования с длительными транзакциями;
- отсутствие ограничений на объем базы данных пространственных данных;
- доступ для пользователей как через локальную сеть, так и через интернет;
- гибкая настройка конфигурации.

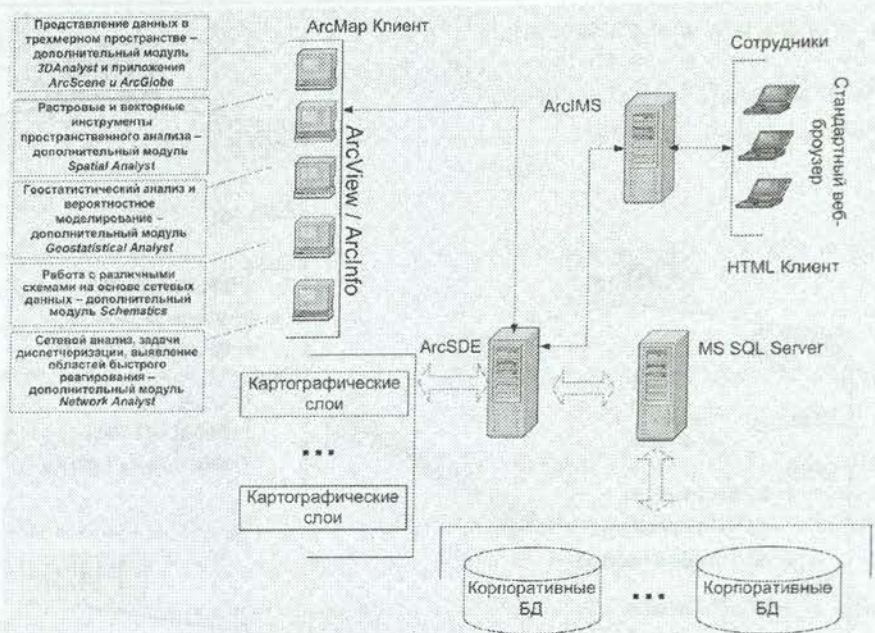


Рис. 5. Структура типового элемента ГИС Росводресурсов

ArcIMS — это масштабируемый картографический Интернет-сервер для публикации карт, данных и метаданных через открытые Интернет-протоколы, который обеспечивает создание ГИС-порталов. Программный продукт ArcIMS позволяет решать следующие задачи:

- публикация карт;
- публикация геоданных;
- публикация метаданных;
- публикация ГИС-функционала;
- публикация ГИС-инструментария.

Модульность ArcGIS 9.x позволяет гибко настраивать функциональность компонент системы. Предполагается, что представленная на рис. 5 архитектура будет являться типовой при построении элементов ГИС Росводресурсов.

Одним из технических методов реализации отношений $F_i^{\text{ОДП}}$ и $F_i^{\text{РМ}}$ в случае однотипности и разнотипности по масштабу элементов является технология распределенного хранения данных.

Распределенная база данных (РБД) — это база данных, расположенная на нескольких удаленных компьютерах, связь между которыми поддерживается соответствующими сетевыми и транспортными протоколами (например, семейством протоколов TCP/IP). Для построения такой базы данных используется технология репликации, под которой понимается возможность копирования данных

из одного источника на множество других и наоборот.

Из теории РБД известно, что РБД должна обладать некоторым набором свойств [6]. Для построения распределенной ГИС Росводресурсов наиболее важными являются следующие:

- **локальная автономия.** Управление данными на каждом из узлов распределенной системы выполняется локально. Это означает, что база геоданных каждого БВУ или территориального отдела функционирует как полноценная локальная база данных; управление ею выполняется локально и независимо от других узлов системы.

- **прозрачная фрагментация и обработка распределенных запросов.** Распределение данных, логически представляющих собой единое целое на различных узлах системы и, соответственно, возможность выполнения операций выборки над распределенной базой данных, сформулированных в рамках обычного запроса, например, на языке SQL.

Фрагментация возможна горизонтальная и вертикальная. Горизонтальная в данном случае означает, например, что фрагменты территории одного масштаба (например, 1:1000000) собираются с территориальных баз геоданных, каждая из которых ведет фрагмент, соответствующий своей территории. Вертикальная фрагментация, например, означает раздельное хранение характеристик объектов, расположенных на нескольких территориях.

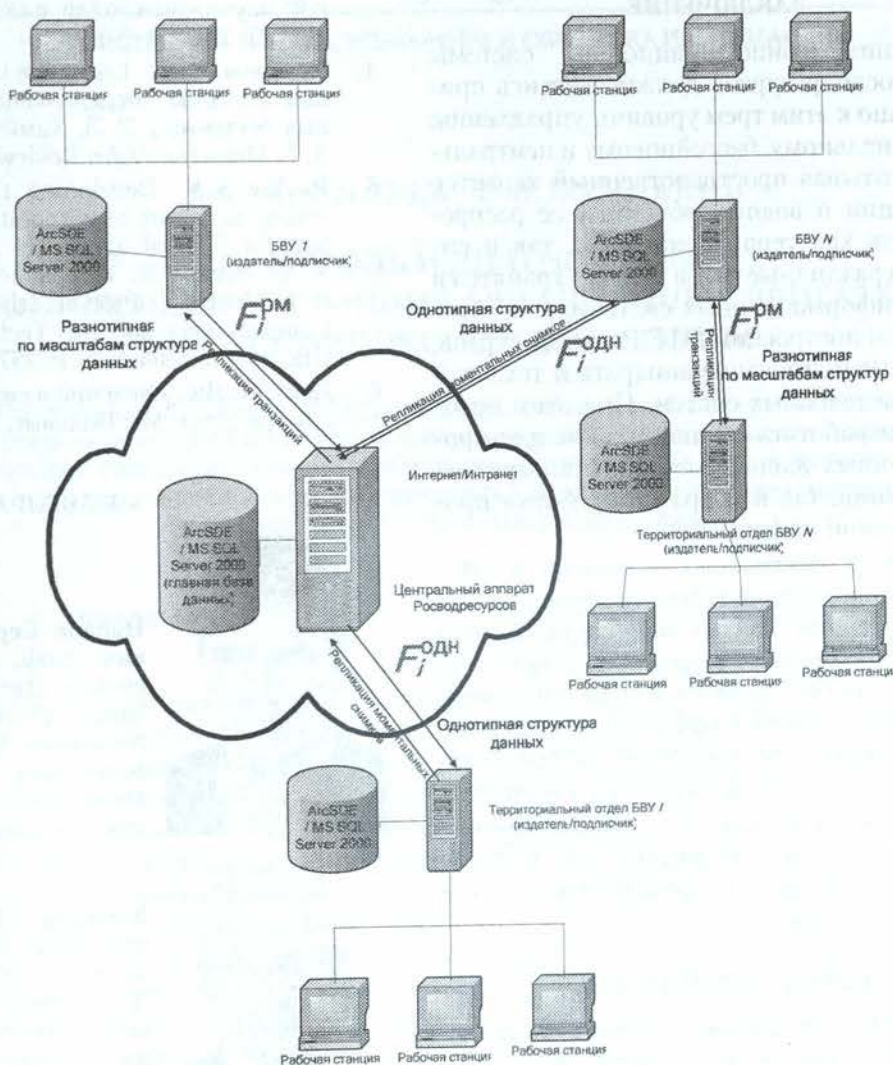


Рис. 6. Некоторые возможные виды реплицирования данных в ГИС Росводресурсов

При организации РБД возникают следующие задачи:

- управление именами в распределенной среде;
- оптимизация распределенных запросов;
- управление распределенными транзакциями.

Решение этих проблем в некоторых случаях при выполнении распределенных запросов приводит к достаточно большим временным затратам на обработку распределенных запросов. Поэтому именно в случае однотипности и разнотипности по масштабу элементов ГИС Росводресурсов возможно организовать взаимодействие на основе механизма репликации данных. Репликация данных — это асинхронный перенос изменений объектов исходной БГД в БГД, принадлежащих различным узлам распределенной системы.

Для построения ГИС Росводресурсов используется программный продукт ArcSDE и СУБД MS SQL Server 2000. Данная архитектура предлагает следующие возможности для реплицирования.

- snapshot replication — репликация моментальных снимков;
- transactional replication — репликация транзакций;
- merge replication — репликация сведений. Наиболее востребованным при построении ГИС Росводресурсов будет являться метод репликации моментальных снимков как метод целостной передачи пространственной и атрибутивной информации между элементами системы в случае нерегламентированного запроса.

Некоторые возможные варианты реализации схемы репликации приведены на рис. 6.

ЗАКЛЮЧЕНИЕ

Создание геоинформационной системы (ГИС) Росводресурсов должно вестись применительно к этим трем уровням управления: территориальному, бассейновому и центральному. Учитывая пространственный характер информации о водных объектах, ее распределенность как территориальную, так и системную (различные части данных хранятся в разных информационных системах), необходимым при построении ГИС Росводресурсов, является использование аппарата и технологий распределенных систем. При этом необходимо разработать специфические для пространственных данных методы и алгоритмы как описания, так и сбора и обработки пространственной информации.

Одним из возможных подходов к описанию распределенных геоинформационных систем является подход на основе расширенных многомерных информационных объектов. На основе данного класса многомерных информационных объектов вводятся отношения типовости элементов распределенной геоинформационной системы и их частей.

Предложенный подход в данный момент применяется при построении 2-й очереди ГИС Росводресурсов на федеральном и территориальном уровнях.

СПИСОК ЛИТЕРАТУРЫ

1. **ВНИИЦ.** Проведение системного анализа информационных потоков в области управления, регулирования и охраны водных ресурсов на федеральном уровне : отчет к гос. контракту В-3-04. Инв. номер ВНИИЦ 02200502893. Уфа : УГАТУ, 2004.
2. **Павлов, С. В.** Интеграция геоинформационных систем с информационными системами трубопроводного предприятия на основе многомерных моделей данных / С. В. Павлов, С. В. Плеханов, Р. И. Бахтизин // Вестник УГАТУ. 2006. Т. 8, № 1 (17). С. 39–42.
3. **Хамитов, Р. З.** Геоинформационная система для информационной поддержки управления водными ресурсами России / Р. З. Хамитов, С. В. Павлов, А. Б. Никитин // Мехатроника, автоматизация, управление : сб. тр. 2-й Все-

рос. научн.-техн. конф. с междунар. участием. 2006. Т. 2. С. 82–87.

4. **Хамитов, Р. З.** Создание геоинформационной системы Федерального агентства водных ресурсов / Р. З. Хамитов, С. В. Павлов, А. Б. Никитин // ArcReview. 2006. № 1 (36).
5. **Pavlov, S. V.** Developing the geoinformation system as a part of corporate information system for Federal agency of water resources / S. V. Pavlov, R. Z. Khamitov, I. G. Ivanov, A. B. Nikitin // Proc. of 8th Int. Workshop on Comp. Science and Inf. Technol. (CSIT'2006). Ufa, Russia, 2006. V. 1. P. 257–261.
6. **Дейт, К. Дж.** Введение в системы баз данных / К. Дж. Дейт. М.: Вильямс, 2005.

ОБ АВТОРАХ



Павлов Сергей Владимирович, проф., зав. каф. геоинф. систем. Дипл. математик по спец. «Вычислительная математика» (БГУ, 1977). Д-р техн. наук (УГАТУ, 1998). Иссл. в обл. обработки пространств. данных.



Хамитов Рустэм Закиевич, рук. Федер. агентства водных ресурсов (ФАВР). Дипл. инж. по авиац. и ракетн. двигателям (МВТУ им. Баумана, 1972). Д-р техн. наук (УГАТУ, 2000). Иссл. в обл. информ. систем при управлении соц.-экон. объектами.



Никитин Алексей Борисович, асс. каф. геоинф. систем. Дипл. инж.-системотехн. (УГАТУ, 2004). Иссл. в обл. геоинф. систем.