

**РАСПРЕДЕЛЕННЫЙ БАНК ГЕОЛОГО-ГЕОФИЗИЧЕСКИХ ДАННЫХ ПО НЕФТИ И
ГАЗУ САХАЛИНСКОЙ ОБЛАСТИ**

В.И. Исаев**, *Н.В. Шпакова, *О.С. Исаева*****

**Томский политехнический университет, г. Томск*

***ГП "Дальневосточный центр информатики и геологических исследований", г. Южно-Сахалинск*

В ГП "Дальинформгеоцентр" на основе интегрированных цифровых пакетов геолого-геофизических данных нефтегазоносных осадочных бассейнов создан распределенный банк данных Сахалинской области. В качестве сервера реляционных баз данных использована СУБД **Oracle**, в клиентской части – графическая система **ArcView**, поддерживающая связь с разнородными СУБД. Локальные вычислительные сети, созданные на основе структурированных кабельных систем, объединены через выделенную некоммутируемую телефонную пару. Скорость передачи данных единой вычислительной сети составляет 5–100 Мбит / сек.

Внутренними средствами **Oracle** и **ArcView** созданы виртуальные представления пользователей и обеспечена прозрачность размещения данных в сети, что решает проблемы избыточности и целостности данных, предельно упрощает работу пользователей с банком данных. Система обновления баз данных "рабочий проект" – "текущий проект" обеспечивает стандартизацию данных, а система администрирования "повсеместной проверки прав доступа" – защиту данных.

Распределенный банк данных является основной информационно-аналитической базой прогнозирования нефтегазоносности осадочных бассейнов и оперативного получения разнообразной информации для управления фондом недр. Основные перспективы развития банка данных заключаются в его наращивании данными по другим осадочным бассейнам Дальневосточного региона и создании региональной корпоративной вычислительной сети.

Ключевые слова: нефтегазоносный осадочный бассейн, банк геолого-геофизических данных, **Oracle**, **ArcView**, вычислительная сеть, система администрирования, прогноз нефтегазоносности, Сахалин, Восток России.

ВВЕДЕНИЕ

В ГП "Дальинформгеоцентр" на основе концептуальной модели [6] с 1995 года осуществляется реализация программы [5] создания Государственного Дальневосточного регионального банка геолого-геофизических данных по нефти и газу. К середине 2001 года модель реализована в виде интегрированных цифровых пакетов по отдельным осадочным бассейнам в варианте несетевого графического интерфейса пользователя [4]. В этой публикации был изложен ряд методологических аспектов проектирования банка данных, представлена информационная и логическая модель, приведены результаты физического проектирования и состояние эксплуатационной готовности баз данных. В работе [3] проблема создания регионального банка геолого-геофизических данных по нефти и газу рассмотрена как одна из проблем оценки нефтегазоматеринского потенциала

осадочных бассейнов Дальневосточного региона. Автоматизированный доступ к информационным ресурсам (ИР) банка данных позволяет оперативно анализировать огромное количество фактического геолого-геофизического материала для решения региональных и поисковых задач.

В настоящей статье рассматривается ряд методологических аспектов и результаты проектирования распределенного [7, 9] банка данных, реализованного на основе ИР нефтегазоносных осадочных бассейнов [1] Сахалина и прилегающего шельфа (рис.1). Будут рассмотрены следующие взаимосвязанные вопросы проектирования распределенного банка данных: создание реляционных таблиц, загрузка данных в СУБД **Oracle** и интеграция всех данных по всем осадочным бассейнам в единую систему; организация сетевой технологии и режима удаленного доступа; администрирование и обновление баз дан-

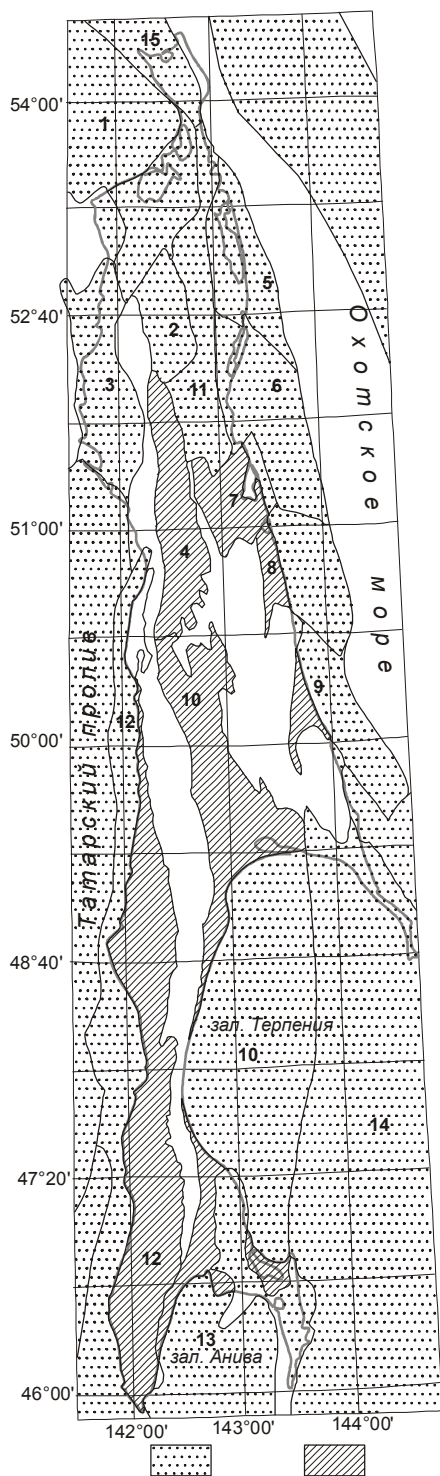


Рис. 1. Обзорная карта размещения осадочных бассейнов Сахалина – объектов формирования и ведения распределенного банка геолого-геофизических данных по нефти и газу (элементы тектонического районирования по В.М. Радюшу, 1988).

1 – площади, границы и номера осадочных бассейнов, прогибов, впадин, поднятий: 1 – Байкальский; 2 – Валский; 3 – Погибинский; 4 – Ныско-Тымский; 5 – Пильтунский (Пильтунская впадина); 6 – Чайвинский (Чайвинская впадина); 7 – Набильский; 8 – Лунский; 9 – Пограничный; 10 – Макаровский; 11 – Дагинский (Дагинское поднятие); 12 – Западно-Сахалинский (Александровский прогиб, Бошняковское поднятие, Ламанонский прогиб, Красногорское поднятие, Чеховский прогиб, Холмское поднятие, Крильонское поднятие); 13 – Анивский, 14 – залива Терпения; 15 – Шмидтовский (Шмидтовское поднятие). 2 – объекты ведения банка данных по состоянию на конец 2001 года.

Понять общую идею сети и распределенных баз данных не составляет труда. А сложные технические детали должны быть "спрятаны" от конечного пользователя. Пользователю предоставляется возможность работать с таким *виртуальным представлением* данных, которое более привычно для него. Обеспечивается *логическая прозрачность данных и прозрачность их размещения*, в результате чего пользователь может формировать запрос ко всем базам, находясь за отдельным клиентским местом, не указывая, куда переслать запрос.*

ЗАГРУЗКА И ОРГАНИЗАЦИЯ ДАННЫХ В СУБД ORACLE

Для создания баз геолого-геофизических данных были использованы системы **ArcView**, **Access**, **Gis DB**, **Internet Explorer**, **Excel** [4, рис. 2]. В качестве СУБД для формирования данных глубоких скважин была задействована **MS Access** [8]. При переходе на клиент-серверную технологию для работы с данными в режиме распределённого банка была привлечена СУБД **Oracle 7** [11]. Переход от **Access** к **Oracle** потребовал изменения в логической структуре баз данных, что позволило решить проблемы *избыточности* и *целостности* данных.

На рисунке 2, в качестве примера, отражён процесс реструктурирования и конвертирования в **Oracle** данных информационных блоков "Керн", "Петрофизика керна", "Химия керна", "Геохимия керна", "Биостратиграфия керна", "Палеонтологические и петрографические коллекции".

При создании общей таблицы "**KERN**" за основу берётся таблица "КАТАЛОГ ИНТЕРВАЛОВ" (информационный блок "Керн") (А). Затем с общей таб-

ных; реализация аналитических запросов пользователей.

Распределенный банк данных очень сложен, особенно в нашем случае, когда банк является неоднородным в отношении используемых СУБД, так как он включает разные типы данных и реализовывался по "восходящей методике проектирования" [4].

* "В то время, как нечто *виртуальное* представляется существующим, но не существует, нечто *прозрачное* представляется несуществующим, но фактически существует" [7].

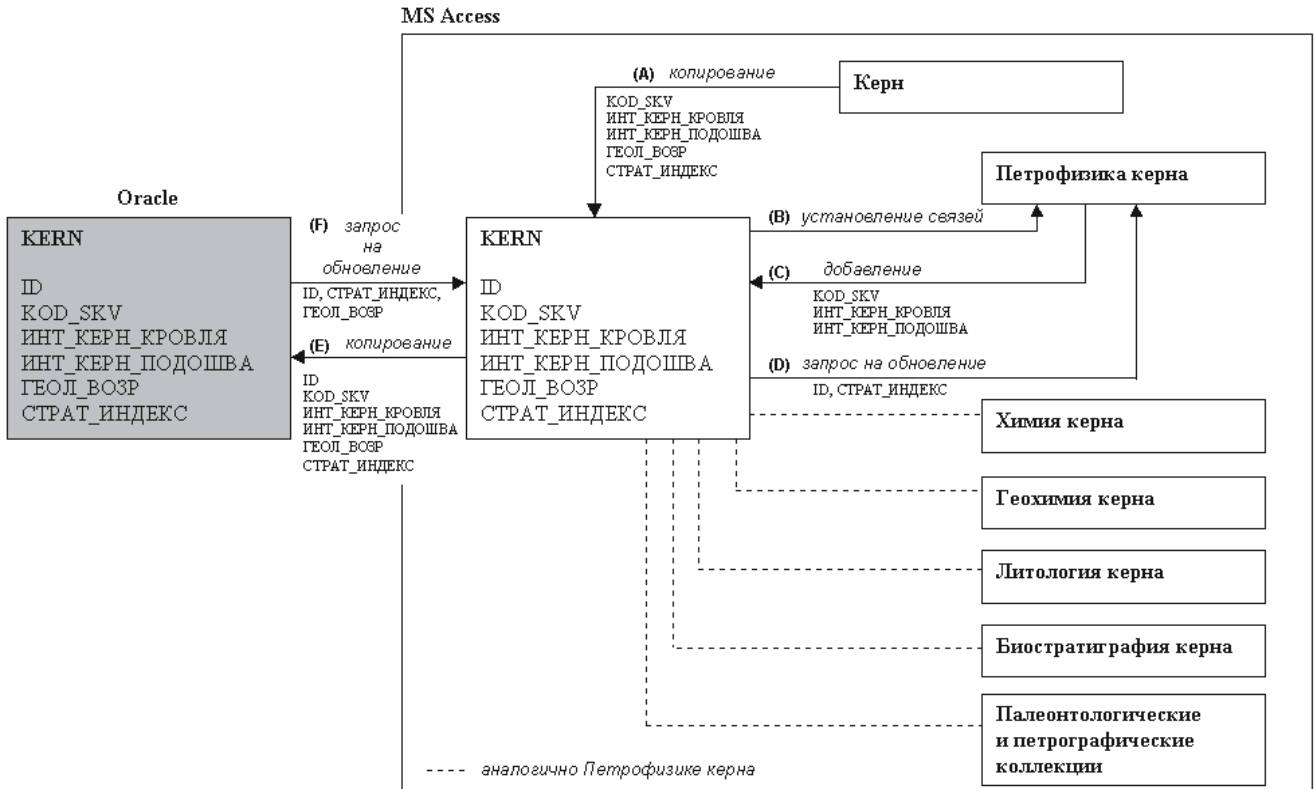


Рис. 2. Процесс реструктурирования и конвертирования из Access в Oracle данных информационных блоков "Керн", "Петрофизика керна", "Химия керна", "Геохимия керна", "Литология керна", "Биостратиграфия керна", "Палеонтологические и петрографические коллекции".

Пояснения в тексте.

лицей устанавливается связь каталогов интервалов вышеперечисленных информационных блоков (B). Если связь не устанавливается по причине отсутствия интервалов в общей таблице, то необходимо добавить недостающие интервалы в общую таблицу (C). Стратиграфический индекс, взятый по "делу скважины" из таблицы "Стратиграфические разбивки", проставляется путём запроса на обновление в вышеупомянутых информационных блоках (D). Если таблица "KERN" в Oracle уже имеет данные, то следует проделать запрос на обновление (F) перед тем, как проделать шаг (D). После того, как будет сформирована общая таблица, данные этой таблицы можно загрузить в Oracle (E).

На рис. 3, в качестве примера, отражен процесс реструктурирования и конвертирования в Oracle данных информационных блоков "Открытый ствол. Опробование на каротажном кабеле" (ОКК) и "Открытый ствол. Гидродинамический каротаж" (ГДК). В поле "INF_BЛОК" проставляется "ГДК" или "ОКК". Для того чтобы исключить дублирующую информацию, каждая строка таблиц идентифициру-

ется индексом ID путём проставления или с помощью запроса на обновление.

На рисунке 4 приведён верхний уровень новой структуры данных в Oracle. А на рисунке 5, в качестве примера, показан логический информационный блок "Испытания в колонне", примыкающий к таблице "Главный каталог глубоких скважин". В новой структуре Oracle таблицы, имеющие одинаковую структуру и взаимосвязи, объединены, а также исключены дублирующие поля.

Во всех таблицах, созданных ранее в Access, убраны поля "расчётных" метаданных. Формирование этих полей в Oracle происходит с помощью разработанных пакетов [12, 13] при запуске "представлений". Так как пересчёт метаданных требует дополнительного времени, то сформированные "представления" записываются в соответствующие таблицы. Просчёт всех баз метаданных (БмД) осуществляется один раз в сутки.

Структура данных в Oracle является прозрачной для пользователя, который работает с виртуальными "представлениями", соответствующими связанной об-

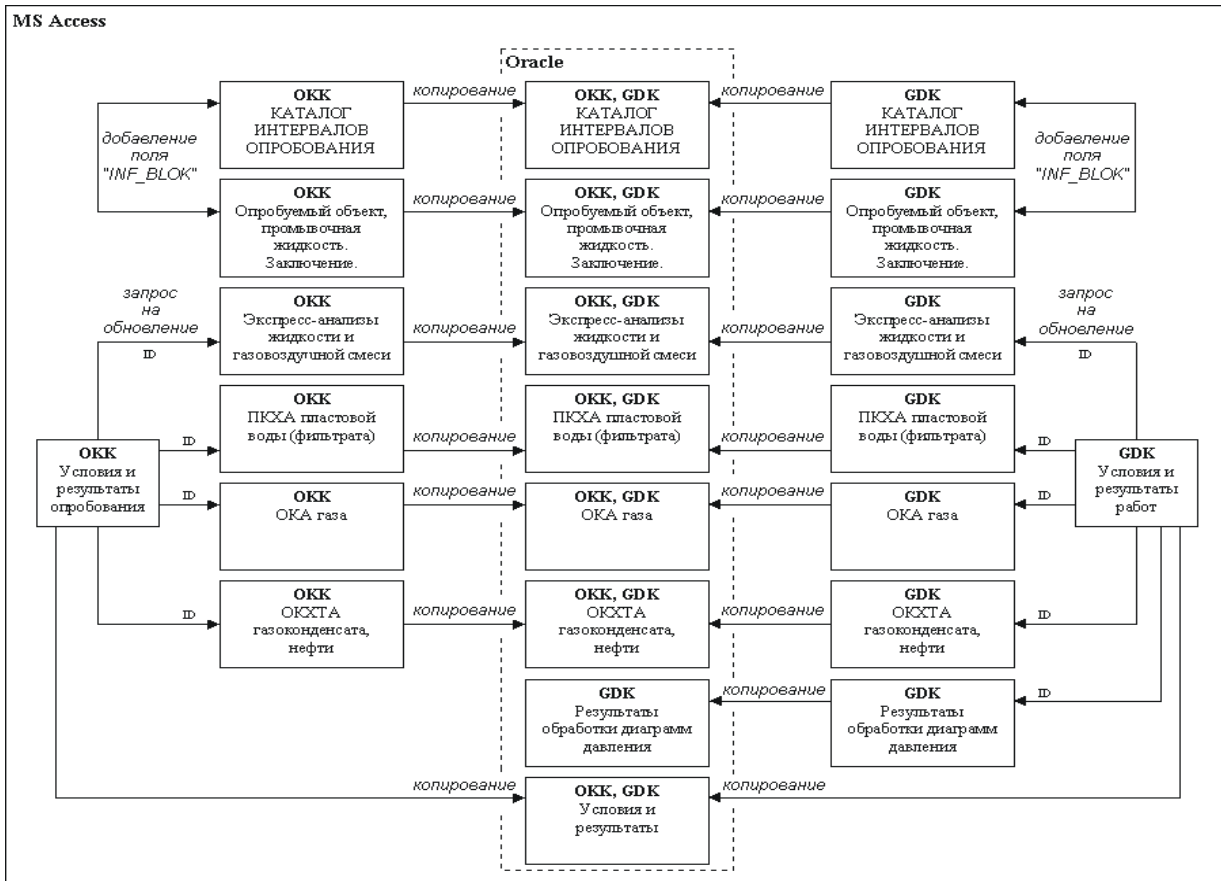


Рис. 3. Процесс реструктурирования и конвертирования из Access в Oracle данных информационных блоков "Открытый ствол. Опробование на каротажном кабеле" (ОКК), "Открытый ствол. Гидродинамический каротаж" (GDK).

ПКХА – полный комплексный химический анализ, ОКА – общий комплексный анализ, ОКХТА – общий комплексный химико-технологический анализ. Остальные пояснения в тексте.

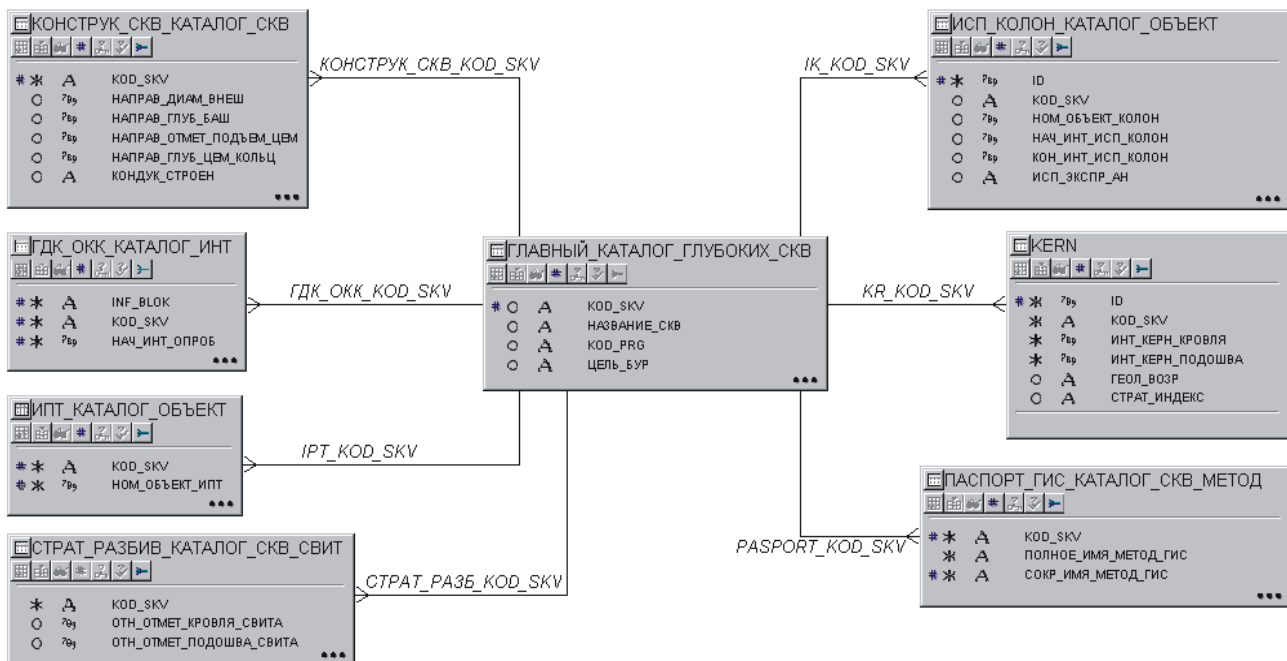


Рис. 4. Верхний уровень структуры данных глубоких скважин в Oracle.

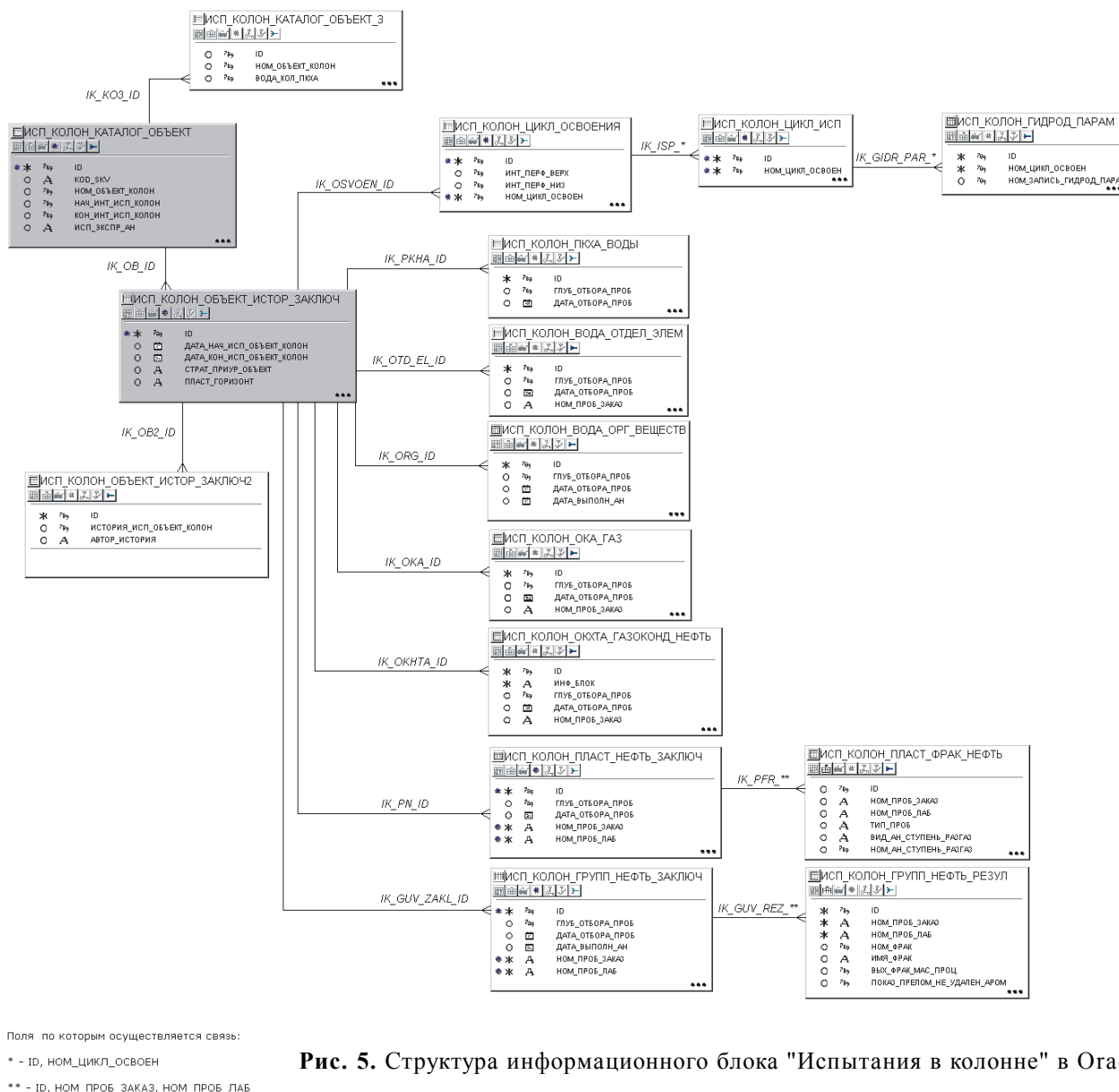


Рис. 5. Структура информационного блока "Испытания в колонне" в Oracle.

щей информационной структуре [4, рис. 2], уже получившей поступательное развитие (рис. 6).

РЕАЛИЗАЦИЯ СЕТЕВОЙ ТЕХНОЛОГИИ И УДАЛЕННОГО ДОСТУПА

В качестве модели реализации распределённого банка данных (многопользовательский режим и удалённый доступ) принята модель *клиент-сервер*, объединяющая удобства персональных компьютеров с централизованными и надёжными серверами (рис. 7). Клиент-серверные системы возлагают интенсивную обработку данных на сервер и оптимизируют сетевой трафик. Системы имеют три компонента: сервер баз данных, клиентское приложение и сеть.

Сервер баз данных (Oracle 7) эффективно и оптимально управляет реляционной базой для множества клиентов, которые одновременно у него этот ресурс запрашивают. Сервер защищает информацию с помощью средств организации доступа к данным, архивации/восстановления, централизованно отслеживает для всех клиентских приложений выполнение правил глобальной целостности данных.

Клиентское приложение – программа *ArcView*, обладающая графическим интерфейсом и возможностью организовывать связь с разнородными СУБД. Клиентское приложение обеспечивает пользователю просмотр данных, выполнение *аналитического* запроса, создание и распечатку отчёта (композиции). В

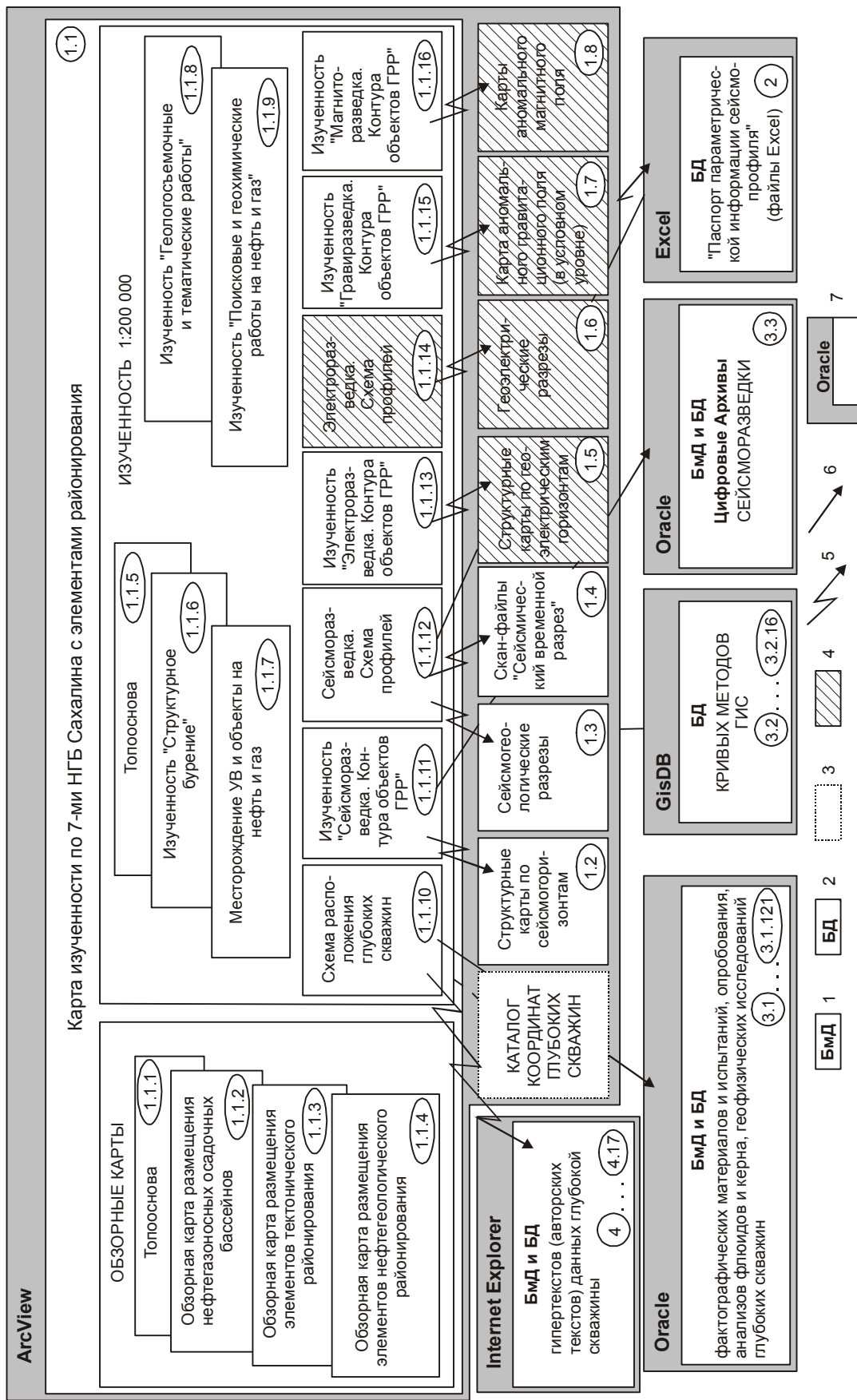


Рис. 6. Общая информационная структура и программная среда (СУБД) распределенного банка геолого-геофизических данных по нефти и газу Сахалинской области.

1 – база метаданных, каталог; 2 – база данных; 3 – информационные материалы на съемных магнитных носителях; 4 – планируемые информационные работы; 5 – “горячая” связь; 6 – связь через SQL-сервер; 7 – программная среда (СУБД).

клиентской части происходит запрос и получение информации от сервера баз данных. Для интерфейса используется встроенный конструктор запросов.

Сетевое программное обеспечение **Oracle SQL*Net** поддерживает все основные сетевые коммуникационные протоколы. Для удалённого доступа к данным **Oracle** с клиентских мест, расположенных как в здании Дальинформгеоцентра, так и за его пределами, используется протокол **TCP/IP**. Локальные вычислительные сети (ЛВС) созданы на базе структурированных кабельных систем с возможностью масштабирования до стандарта **100 Base-T**. Компьютеры в настоящее время оснащены сетевыми картами со спецификацией **10–100 Base-T**. Объединение ЛВС Дальинформгеоцентра и ЛВС Комитета природных ресурсов (КПР) по Сахалинской области, расположенных на расстоянии порядка **500 м**, выполнено на основе выделенной некоммутируемой телефонной пары с помощью модемов **Pair Gain MM-300S**. Скорость передачи данных единой вычислительной сети (ЕВС) составляет **5–100 Мбит/сек**.

Для удалённой работы с задачами, которые не являются первоочередными для перевода на технологию клиент-сервер, используется технология файл-сервера.

Пользовательский проект банка данных ("рабочий проект") представляет собой модуль [10] взаимосвязанных, но разделенных по узлам сети графических данных и табличных баз данных и включает в себя (рис. 8):

- 1) базу данных цифровых картографических материалов, сформированную в форматах ArcView;
- 2) базу данных полнотекстовых описаний по глубоким скважинам, сформированную в формате HTML;
- 3) базу каротажных кривых, сформированную в форматах Gis DB;
- 4) базу паспортов параметрической информации профилей сейсморазведки, сформированную в формате Excel;
- 5) базу метаданных и данных по глубоким скважинам, сформированную в программе Oracle;

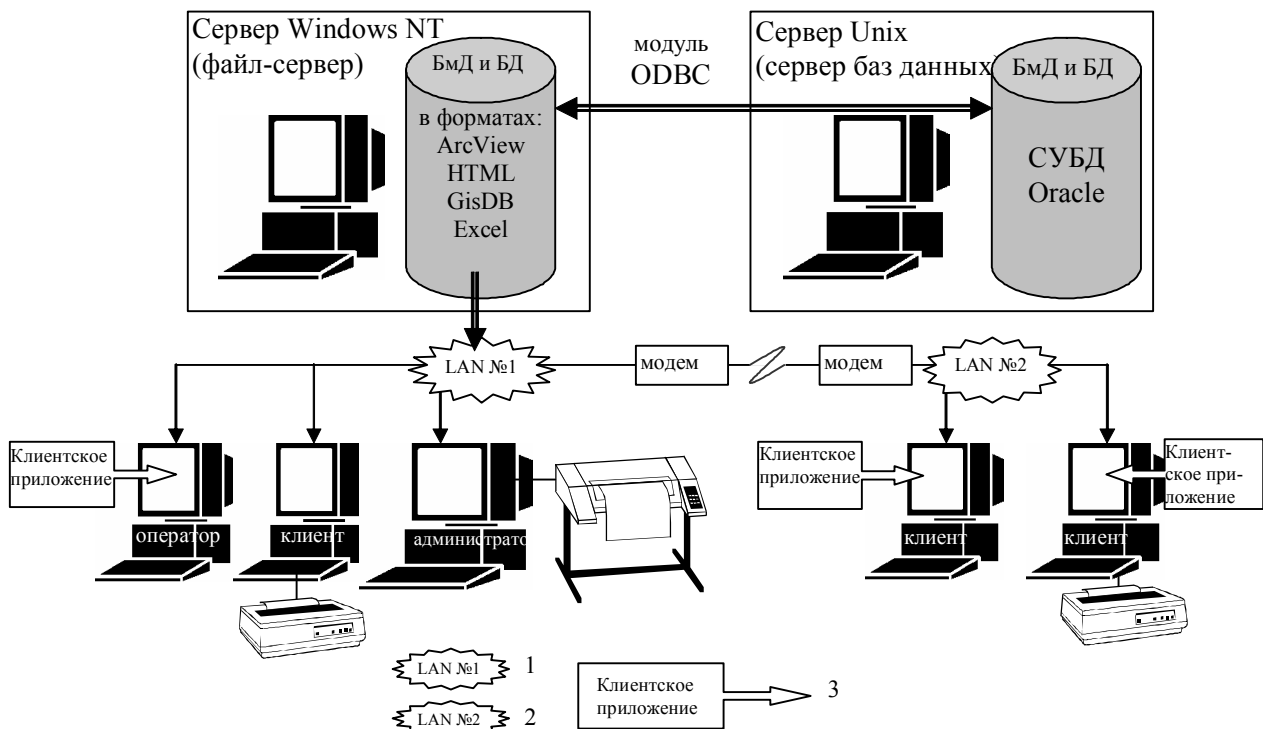


Рис. 7. Модель реализации распределенного банка геолого-геофизических данных по нефти и газу Сахалинской области.

1 – локальная вычислительная сеть Дальинформгеоцентра; 2 – локальная вычислительная сеть КПР по Сахалинской области; 3 – клиентское приложение, включающие программы ArcView, Oracle client, GisDB, Excel, Internet Explorer.

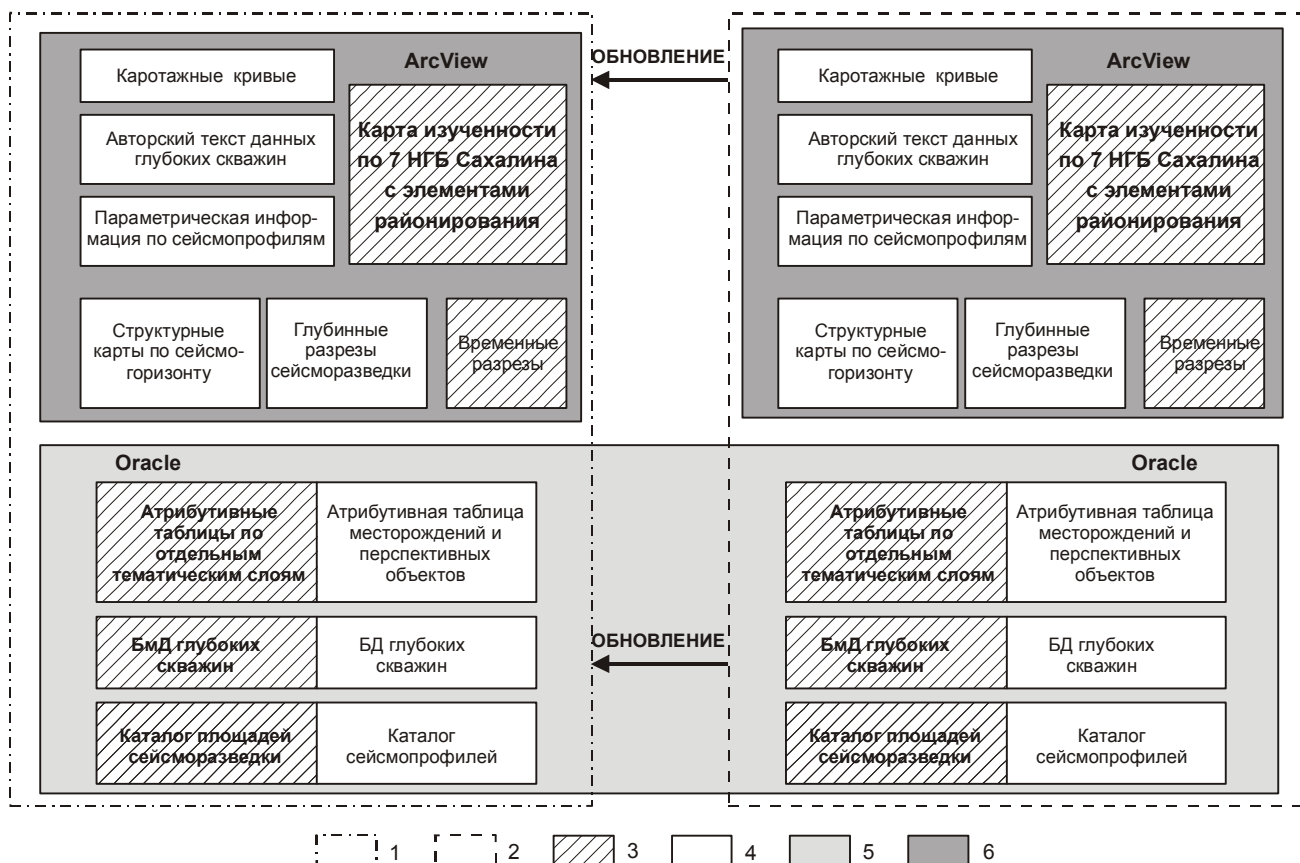


Рис. 8. Схема размещения, администрирования и обновления “рабочего проекта” банка данных.

1 – “рабочий проект”, 2 – “текущий проект”, 3 – общедоступные данные, 4 – данные с регламентированным доступом, 5 – данные, расположенные на сервере баз данных, 6 – данные, расположенные на файл-сервере.

б) базу справочно-ссылочных данных сейсморазведки, сформированную в программе Oracle.

Наличие единой концептуальной схемы всей сети (рис. 6, 7) и схемы, определяющей местонахождение данных в сети, “прописанной” внутренними средствами ArcView (скрипты ArcView), атрибутивными таблицами в Oracle и поддерживаемой посредством связи с SQL-сервером, обеспечивает логическую прозрачность данных и прозрачность размещения данных. В результате чего пользователь может формулировать запрос ко всем базам, находясь на отдельном клиентском месте, при этом не указывая, куда переслать запрос, чтобы получить требуемые данные. То есть пользователь как бы работает с централизованным банком данных.

АДМИНИСТРИРОВАНИЕ И ОБНОВЛЕНИЕ БАЗ ДАННЫХ

Здесь администрирование рассматривается в узком смысле – в обеспечении “секретности” в сети, что достигается методом повсеместной проверки прав доступа пользователей.

Администрирование банка данных условно делится на две части. Администрирование картографических данных, расположенных на файл-сервере, производится с помощью операционной системы Windows NT 4.0. Администрирование табличных данных, расположенных на сервере баз данных, осуществляется средствами СУБД Oracle 7.

Все пользователи, имеющие доступ к ЛВС или ЕВС, имеют свободный доступ для просмотра общедоступных данных “рабочего проекта” (рис. 8). Доступ к данным с регламентированным доступом определяется реестром пользователей. При составлении реестра предполагается, что пользователь обращается только к тем частям баз данных “рабочего проекта”, которые необходимы ему для работы. В реестре учитывается: фамилия, должность пользователя; организация и отдел, где работает пользователь; место расположения компьютера пользователя; IP – адрес; список ограничений на картографические данные; список ограничений на табличные данные. Реестр составляется администратором банка данных на основе

заявок начальников отделов Дальинформгеоцентра и КПР, согласовывается с руководителем геологической службы КПР и утверждается директором Дальинформгеоцентра.

Внешнему пользователю "рабочего проекта" предоставляется право "только чтения", правом "обновления" обладает только администратор банка данных.

Для обновления находящегося в эксплуатации банка данных – "рабочего проекта" – формируется "текущий проект", к которому имеет доступ, персонал, формирующий банк данных (рис. 9).

"Рабочий проект" периодически обновляется на "текущий" администратором банка данных на файл-сервере и на сервере баз данных (рис.8). Резервное копирование банка данных, расположенного на файл-сервере и сервере баз данных, осуществляется администратором сети.

АНАЛИТИЧЕСКИЕ ЗАПРОСЫ ПОЛЬЗОВАТЕЛЕЙ

На конец 2001 года распределенный банк данных Сахалинской области состоит из цифровой информации объемом 143950 Мбайт, из них – 140155 Мбайт составляет архив сейсмозаписей, на которые имеются адресные ссылки.

В настоящее время ИР банка данных интенсивно используются в автоматизированном режиме в научных и тематических исследованиях осадочных бассейнов региона с целью оценки их нефтегазового потенциала, а также с целью объемно-площадного изучения общих закономерностей степени катагенеза и литоплотностных характеристик мезокайнозойских толщ и свит [2]. Основной информационной базой таких исследований являются цифровые глубинные и временные разрезы сейсморазведки, структурные карты, цифровые базы кривых каротажа, а также цифровые базы данных глубоких скважин (данные о

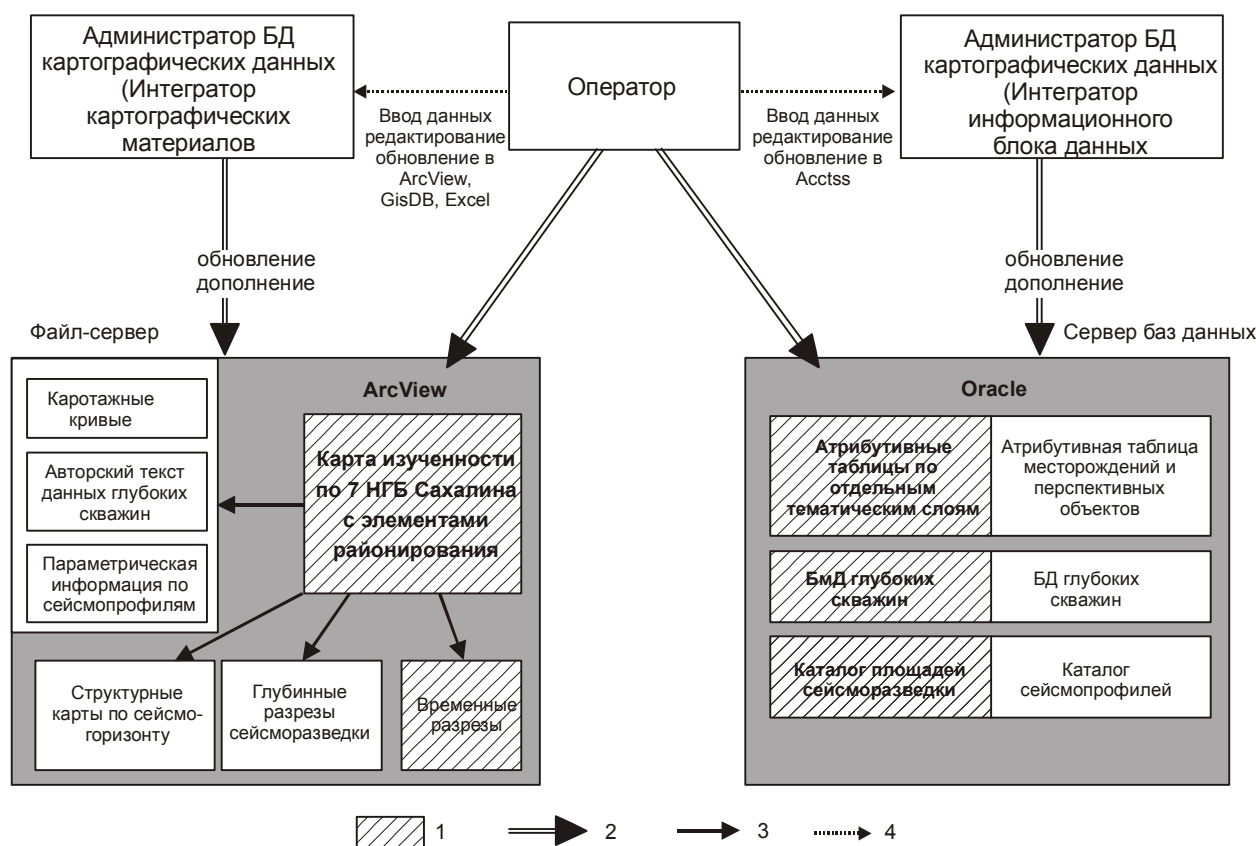


Рис. 9. Схема размещения, администрирования, обновления и дополнения "текущего проекта" банка данных.

1 – общедоступные данные, 2 – регламентированный доступ, 3 – регламентированный доступ по "горячей" связи, 4 – передача данных.

конструкции, послойное описание керна, петрофизика и литология керна, стратиграфические разбивки). Новейшие результаты таких исследований сами, в свою очередь, создают компьютерную информационную базу объемно-площадного распределения подсчетных параметров для оценки (переоценки) прогнозных ресурсов нефти, газа и конденсата [3].

Для сопоставления результатов математического моделирования и прогнозирования с прямыми признаками нефтегазоносности и палеообстановками осадконакопления используются цифровые карты месторождений и объектов на нефть и газ с атрибутивными таблицами по стратиграфической приуроченности, геометризации объекта, запасам и ресурсам, а также используются цифровые базы глубоких скважин (данные испытаний, опробования и анализов флюидов, данные по биостратиграфическим анализам).

Распределенный банк данных предоставляет возможность практически любых аналитических запросов как в процессе научных и тематических исследований, так и для оперативного получения разнообразной геолого-геофизической информации для нужд управления фондом недр.

ЗАКЛЮЧЕНИЕ

1. Научно-исследовательскими и проектными работами, выполняемыми в ГП «Дальинформгеоцентр», решается одна из научно-технических проблем оценки нефтегазового потенциала осадочных бассейнов Дальневосточного региона – создается не имеющий аналогов на Востоке России распределенный банк геолого-геофизических данных по нефти и газу.

2. Для реализации банка данных привлечены и освоены современные, относительно недорогие промышленные системы **Oracle** и **ArcView**, позволившие организовать эффективную клиент-серверную технологию обработки данных в режиме удаленного доступа.

3. Практически воплощены основные методологические принципы проектирования – *виртуальность* представлений пользователей и *прозрачность* размещения данных, что решает проблемы избыточности и целостности данных, а также предельно упрощает работу пользователя с банком данных.

4. Спроектирована и действует система обновления и администрирования баз данных, обеспечивающая стандартизацию данных и их защиту.

5. Распределенный банк данных Сахалинской области стал основной информационно-аналитичес-

кой базой научных и тематических исследований нефтегазоносности осадочных бассейнов, охваченных банком данных, и оперативного получения разнообразной геолого-геофизической информации при управлении фондом недр.

6. Перспективы развития распределенного банка данных заключаются в наращивании его информационной базы данными по другим осадочным бассейнам Дальневосточного региона и предоставлении информационных услуг клиентам всего региона на основе корпоративной вычислительной сети.

В проектировании распределенного банка данных по нефти и газу помимо авторов участвуют специалисты Дальинформгеоцентра Д.В. Москаленко, А.А. Юрчук, А.Т. Ким, Р.С. Ляшкевич, А.В. Бакунов. Работа по переводу баз данных на платформу СУБД **Oracle** консультировали специалисты Центра информационных технологий ОАО «Роснефть – Сахалинморнефтегаз» Г.Н. Жукова и М.Г. Жуков. Авторы весьма признательны Е.А. Суслу за помощь, оказанную при подготовке статьи к печати.

ЛИТЕРАТУРА

1. Варнавский В.Г., Коблов Э.Г., Буценко Р.Л., Деревскова Н.А., Иваньшина Л.П., Кириллова Г.Л., Крапивенцева В.В., Кузнецов В.Е., Тронова Т.Ю., Уткина А.И. Литолого-петрофизические критерии нефтегазоносности. М.: Наука, 1990. 270 с.
2. Исаев В.И. Прогноз материнских толщ и зон нефтегазонакопления по результатам геоплотностного и палеотемпературного моделирования // Геофиз. журн. 2002. № 2. С. 60–70.
3. Исаев В.И., Косыгин В.Ю., Соловейчик Ю.Г., Юрчук А.А., Гуленок Р.Ю., Шпакова Н.В. Проблемы оценки нефтегазоматеринского потенциала осадочных бассейнов Дальневосточного региона // Геофиз. журн. 2002. №1. С.28–52.
4. Исаев В.И., Юрчук А.А., Шпакова Н.В., Войкова С.И., Исаева О.С., Соколова В.В. Государственный Дальневосточный региональный банк геолого-геофизических данных по нефти и газу (модель реализации) // Тихоокеан. геология. 2002. №1. С.111–126.
5. Кисловский О.А., Исаев В.И. Проект работ по созданию Дальневосточного регионального банка цифровой геолого – геофизической информации по нефти и газу (ДВ АбнД-НГ). М.: Роскомнедра, 1995. 27 с.
6. Кисловский О.А., Исаев В.И. О концептуальной модели создания Государственного Дальневосточного регионального банка геолого-геофизической информации по нефти и газу (ДВ АбнД-НГ) // Тихоокеан. геология. 1998. №1. С. 131–140.
7. Мартин Дж. Вычислительные сети и распределенная обработка данных: программное обеспечение, методы и архитектура. Вып. 1. М.: Финансы и статистика, 1985. 256 с.
8. Нортон П., Андерсен В. Разработка приложений в Access 97. СПб.: BHV-Санкт-Петербург, 1998. 656 с.

9. Озкарахан Э. Машины баз данных и управление базами данных. М.: Мир, 1989. 696 с.
10. Проектирование корпоративных ГИС. Организация баз данных. ESRI-М.: Проектный и Учебно-методический Центр по геоинформационным технологиям "ГИС-проект", 1999. 807 с.
11. Austin D., Caindec L., Frazier J. и др. Сервер Oracle 7.3: Администрирование баз данных. Учеб. пособие. Ч. 1. Oracle Corporation, 1996. 459 с.
12. Collins J. Непроцедурное программирование на языке SQL в среде SQL* Plus. Oracle Corporation, 1994. 786 с.
13. Collins J. Процедурное программирование в Oracle с использованием PL/ SQL. Oracle Corporation, 1994. 350 с.

Поступила в редакцию 10 июня 2002 г.

Рекомендована к печати Б.И. Бурде

V.I. Isaev, N.V. Shpakova, O.S. Isaeva

The distributed bank of the geological and geophysical data on the oil and gas of the Sakhalin area

In the "Dalinformgeocenter", on the basis of the integrated digital packets of the geological and geophysical data of oil-and-gas sedimentary basins, the distributed data bank of the Sakhalin area has been created. As the server of relational databases, the graphic system **ArcView**, communicating with heterogeneous DBMS has been used by the DBMS **Oracle** in the client part. The local computer networks created on the basis of structured cable systems are united through an isolated non-switched telephone pair. The transfer rate of the unified computer network makes 5-100 Mbit / sec.

The internal means of the **Oracle** and **ArcView** create virtual representations of the users, and the transparency of arrangement of the data in the network is ensured, which solves the problems of redundancy and data integrity, and extremely simplifies the activity of the users with the data bank. The system provides updating of the databases «work project» - «current project», standardizes data, and the system of management of "all-round verification of access rights" ensures data protection.

The distributed data bank is the basic informational-and-analytical base of prediction of oil and gas potential of sedimentary basins and operational obtaining of miscellaneous information for controlling the entrails fund. The basic prospects for the development of the data bank involve piling up data of other sedimentary basins of the Far East region and creation the regional corporate computer network.