

Информационные технологии и управление данными

Seabed – дальнейшее развитие баз данных для разведки и эксплуатации

Seabed marks continuing evolution of the E&P database

Питер Уэбб (Peter Webb) и Дин Куигли (Dean Quigley) из компании Schlumberger Information Solutions разъясняют вопросы создания новых баз данных для применения в разведке и эксплуатации на примере своей цифровой модели данных Seabed.

Исторически сложилось так, что создание баз данных велось в условиях ограничений по технике и по задачам. Стоимость и производительность операционных систем, сетей и оборудования устанавливали пределы сложности строения и объемам данных которые можно хранить. Но сегодня вполне доступная новая техника имеет большую производительность и возможности по хранению данных, поэтому пора применять более жесткие подходы к созданию баз данных (БД).

Обзор баз данных

Десять лет назад компания Schlumberger Information Solutions (SIS), проведя исследование использования и устройства БД, выявила четыре основных типа БД и установила их отличия (табл. 1):

Главное хранилище – Известное также под названием «корпоративная БД», главное хранилище является единственным местом, где доступны проверенные данные и информация. Каждый элемент данных может существовать в нескольких версиях, но правильная, самая свежая версия находится в главном хранилище.

Главное хранилище не обязательно является единой монолитной БД; иногда оно может формироваться из нескольких связанных или подчиненных хранилищ. Основной функцией главного хранилища является предоставление доступа к самым свежим достоверным и проверенным данным.

БД проекта – Основной функцией БД проекта является предоставление ряду программ, рабочих групп и сотрудников совместного доступа к данным и результатам. В БД проекта обычно находятся новейшие версии результатов интерпретации. В большинстве систем интерпретации БД проекта также содержит рабочую копию данных. Так называемая «оперативная БД» является разновидностью БД проекта, сделанная для специфических целей, например частого переключения между потоками данных.

Архивная БД – Целью архивной БД является хранение в течение неопределенно долгого времени всех версий наборов данных, результатов интерпретации и отчетов, в том числе исходных полевых данных. Данные снабжаются меткой даты, времени и владельца, поэтому каждый может проследить последовательность внесения изменений.

БД-каталог – Главный каталог работает с метаданными, т. е. с данными о данных в каждом хранилище, будь то БД проекта, главное хранилище или архивная БД.

В исследовании SIS показано, что необходимы все четыре типа БД, и между БД всех типов должна существовать беспрепятственная связь.

Одинаковый размер всем не подходит

Одним из мифов о БД для разведки и эксплуатации является мнение, что всем нужны БД одного и того же размера, поэтому создаются огромные монолитные БД, работающие со всеми данными предприятия. Даже лучшие современные методы поиска не работают в очень больших БД также быстро, как в БД меньшего размера. Низкая скорость доступа заставила интерпретаторов отказаться от использования таких БД.

Хранение записей

При некоторых операциях все данные стараются хранить на рабочих станциях, а для отслеживания содержания и состояния каждого проекта хранят записи соответствующего содержания. Такой подход может работать для нескольких проектов на начальной стадии, но обычно кончается огромным складом дисков или ящиков с лентой, которые никто никогда не читает по второму разу. В результате

Свойство	БД проекта	Главная БД	Архивная БД	БД-каталог
Безопасность	Чтение и запись участниками проекта	Допуск к чтению уполномоченных лиц, к записи - администратора	Допуск к чтению уполномоченных лиц, запись - в коде внесения в архив	Допуск к чтению уполномоченных лиц, запись - при регистрации.
Функции	Управление данными при анализе и интерпретации, Доступ для программ	Управление проверенными достоверными данными. Поиск, запросы и отбор данных проектов	Долгосрочное хранение данных и результатов, внесенных в каталог	Хранение списков элементов данных и мест хранения. Просмотр таких мест хранения
Типы данных	Цифровые рабочие копии данных с мгновенным (on-line) доступом	Цифровые и физические проверенные данные. Разные формы доступа (on-line, off-line)	Цифровые и физические данные по завершенным проектам, файлы в исходных форматах. Данные, которые следует хранить	Список всех элементов данных и мест их хранения. Метаданные
Скорость	Оптимизирована для параллельного доступа многих пользователей	Оптимизирована для поиска и извлечения больших объемов	Оптимизирована для поиска объектов архива	Оптимизирована для эффективной регистрации мест хранения и работы с ними
Размер	От малого, до среднего. Сотни элементов данных, иногда - больших.	От малого до очень большого. Миллионы элементов данных, иногда - больших.	От среднего до очень большого. Сотни больших элементов данных,	Малый (список). Миллионы малых элементов

Табл. 1. Типы БД, применяемые в разведке и эксплуатации

Информационные технологии и управление данными

может произойти интерпретация не тех данных и даже повторение ошибок, поскольку выученные уроки пылятся в коробках на складе. Через несколько лет работы с таким подходом понадобятся огромные усилия, чтобы отобрать и проверить нужные данные.

Различные модели данных

Следующим моментом в разработке оптимальной БД для разведки и эксплуатации является порядок работы с информацией в моделях данных, принятых в отрасли. Многие разработчики программ уделяют внимания только требованиям какой-то одной сферы, ориентируясь на характерные для нее данные. Существуют модели данных бурения, каротажа, сейсморазведки, добычи. Программы оптимизируются для работы с данными и хранения данных в форматах, удобных для этих специфических областей.

Разработка программ никогда не велась в едином, логичном для всего спектра задач отрасли, ключе, что позволило бы соответствовать как общим целям, так и специфическим требованиям той или иной области. Главными проблемами являются необходимость работать с большим количеством типов данных и необходимость представлять различные схемы работы.

В последние годы наибольшие успехи в применении БД в разведке и эксплуатации достигнуты при создании хранилищ, должным образом разделенных для разных целей. Такие операции поддерживаются главным каталогом (БД-каталог). Но БД-мечта, связывающая главное хранилище, БД проектов, архивную БД и БД-каталог до сих пор не появилась.

БД нового поколения

Компания Schlumberger разработала новую модель данных и новую систему управления БД (СУБД) для предприятий развивающейся отрасли разведки и эксплуатации на основе технологии *off-the-shelf*. Эта проблемно-ориентированная БД обеспечивает гибкое и эффективное осуществление требований бизнеса и нужд практической работы в нефтегазовой отрасли. Проще говоря, проект Seabed дает единую модель управления данными в масштабе корпорации, проекта, текущей деятельности и в реальном времени, которую можно разбить на подмножества в виде отдельных модулей по сфере деятельности, по функциям и по степени детальности. Проект не является отдельным коммерческим продуктом или заменой программы Finder или другой оболочки для работы с информацией, но представляет собой полноценное средство хранения со средствами поддержки. Это развивающаяся программа модели БД, основанная на идее долгосрочного подхода к управлению информацией для будущих нужд отрасли.

Модель Seabed является моделью данных и СУБД общего назначения, снабженной средствами обслуживания БД и родственных программ, доля поддержки программ, систем и процедур, созданных SIS. Логическая модель данных отвечает требованиям всех четырех упомянутых типов БД и разбита на модули по сферам деятельности. Ее устройство допускает разделение

на отдельные блоки по сферам деятельности, степени детальности и по функциям, избегая тем самым подхода «один размер подходит всем».

Работая на серверах Oracle и SQL, в операционных системах Solaris, Linux и Windows, модель Seabed можно масштабировать от однопользовательского уровня до уровня предприятия. Например, в семействе программ для нужд эксплуатации Avocet важные части модели данных можно «отделить» и установить на рабочем месте, оптимизировав реализацию под конкретную сферу деятельности и способ использования. Модель данных по-прежнему соответствует Seabed, и, значит, данные, хранящиеся на рабочем месте, можно легко перенести и интегрировать в любую другую программу, использующую модель Seabed в сфере добычи.

Опыт показал, что для нужд отрасли в целом требуются жесткие, закрепленные в модели данных стандарты, подобные стандарту POSC (Petrotechnical Open Standards Corporation), а для повседневной работы удобнее модели, подобные PPDM (Public Petroleum Data Model). В модели Seabed использованы лучшие стороны обоих, а также возможность ввода из ряда других источников - OpenSpirit, Wellsite Information Transfer Standard Markup Language (WITSML), American Petroleum Institute Recommended Practice 66 (RP66) и Oilfield Services Data Dictionary (OSDD).

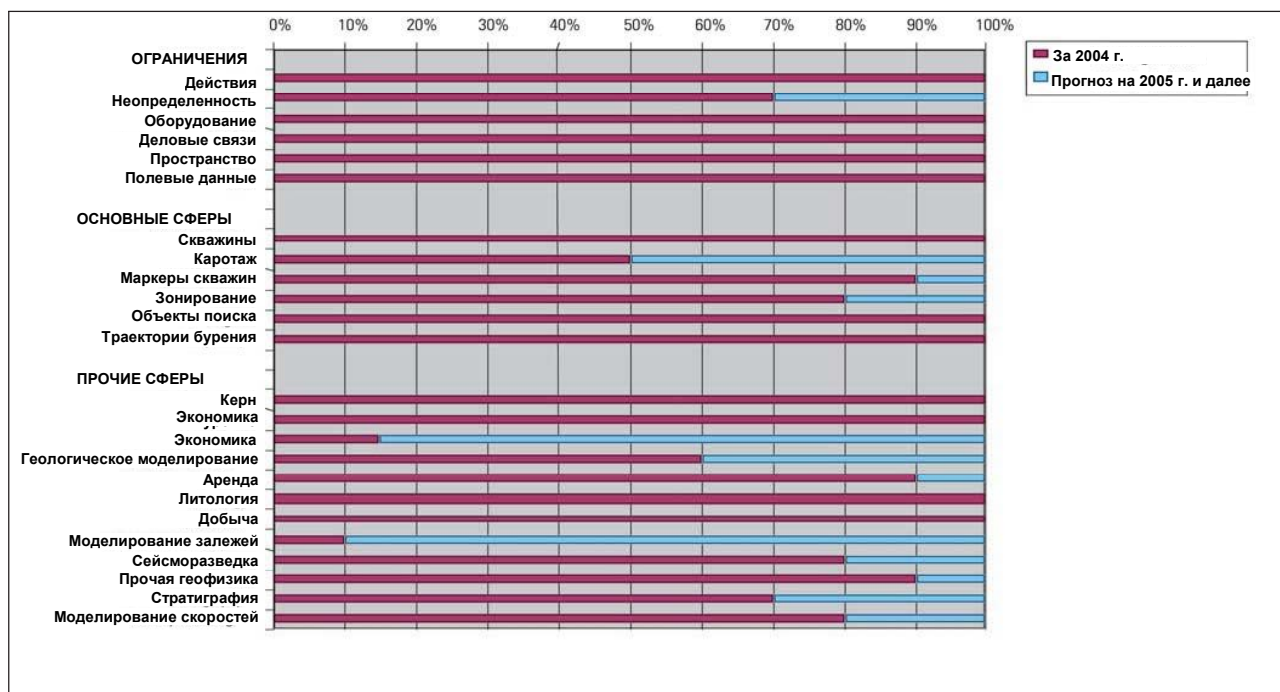
Простота доступа

Средства загрузки и выгрузки (экспорта данных) сочленяются с программами, основанными на модели Seabed, через программы-адаптеры. Средства загрузки/экспорта форматов ASCII, электронных таблиц Spreadsheet, Каротажного стандарта ASCII (LAS) и стандарта Британской ассоциации морских работ (p1/90) резко улучшают загрузку и экспорт данных, доступ, поиск, формирование запросов и передачу данных между рядом

Служебные функции БД	
Безопасность и доступ	На уровне ролей для серверов Oracle и SQL, на уровне очереди для Oracle с использованием OLS
Преобразование модулей	Автоматическое преобразование через настроенные интерфейсы
События с данными	Регистрация запросов на обновление, добавление, и/или удаление данных
Маскировка реализации	Абстрагирование таблиц БД и интерфейсов чтобы пользоваться новыми свойствами не затрагивая пользователей и программы
Сравнение и слияние	Изменяемые правила сравнения/слияния при загрузке перед выполнением
Расширяемость	Настройки на уровне доступа к данным. Средства сравнения структур БД
Обновление проекта	Средства обновления версий, включая безопасный перенос настроек
Прикладные функции БД	
Управление слиянием	Изменяемые правила сравнения/слияния при загрузке
Управление БД	Интерфейс администратора, включая управление пространством и списками, расширением, обновлением и безопасностью
Загрузка/экспорт ASCII	Поддерживает различные форматы и управляющие файлы
Загрузка/экспорт электронных таблиц	Поддерживает любые ASCII-таблицы, включая Excel и CSV
Загрузка/экспорт LAS	Поддерживает LAS 2.0
Загрузка/экспорт UKOOA	Поддерживает UKOOA p1/90. Управляющие файлы можно редактировать для других форматов
Проверка полевых данных	Поддерживает средства просмотра полевых данных, например, диаграмм каротажа

Табл. 2.

Информационные технологии и управление данными



программ и сфер деятельности удобным и эффективным образом. Эти операции были узким местом процессов работы и работы с данными во всей отрасли.

Функции разделены на служебные (Services) и прикладные (Applications). Служебные функции встроены в физическую реализацию СУБД или тесно с ней связаны. Прикладные функции реализованы отдельными программами, созданными в связи с БД (табл. 2).

Защита данных и информации

Защита данных всегда была приоритетной для БД, а в последнее время стала особенно важна из-за участвовавших злонамеренных вторжений. Кроме того, гарантии конфиденциальности путем упорядочения доступа очень важны, когда к одним и тем же данным обращаются разные компании. Раньше полную безопасность можно было обеспечить, только создавая отдельную БД на каждый такой запрос, но за такое

Решение приходилось окупать большими трудозатратами. Альтернатива состояла в том, чтобы регулировать доступ с помощью внешней по отношению к БД программы, но этот упрощенный подход зависит от среднего уровня, где предоставляются привилегии.

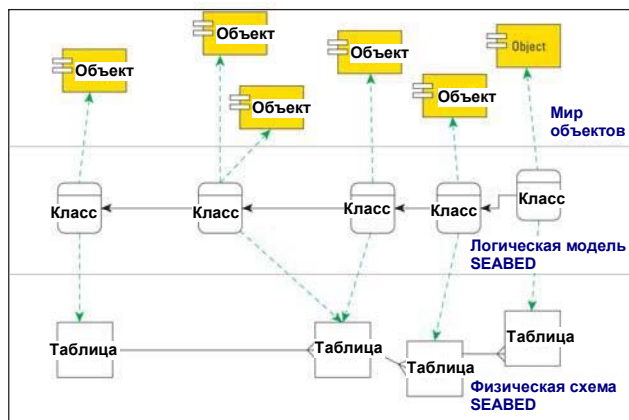
В модели Seabed используются системы безопасности серверов Oracle и SQL, обеспечивающие контроль данных с помощью механизма Ролей (Roles). Реализации Oracle, кроме того, обеспечивают безопасность на уровне очереди (row), используя аппарат Oracle Label Security (OLS). Таким образом, СУБД сама решает вопрос упорядочения доступа и не зависит от внешних программ, созданных третьими сторонами.

Управление БД

Администрирование БД – дело трудоемкое. Помимо повседневных процедур, есть крупные задачи – установка программ, обновление версий, внутренний и внешний перенос данных, настройка рабочих мест – которые следует тщательно планировать и выполнять так, чтобы уменьшить потери и риски для бизнеса.

На стадии разработки модели Seabed были решены несколько важных вопросов:

- Быстродействие
- Быстрый доступ к данным и средства, обеспечивающие стабильность цикла бизнеса.
- Расширяемость
- Поддерживается настройка на местах под конкретные затруднения и варианты действий.
- Настройки безопасным образом сохраняются при обновлениях, что ускоряет освоение новых версий.
- Масштабируемость
- Число записей – от одной до сотен миллионов.
- Число рабочих мест – от одного до всего предприятия.



Информационные технологии и управление данными

- Гибкость конфигурации
- Модульное устройство
- Главная программа совместима с инструментами текущего процесса.
- Многоплатформенность
- Oracle, SQLServer
- Windows, Solaris, Linux

Может возникнуть необходимость локальной настройки модели данных к типам данных, недоступных в настоящее время для совместимости с внутренними системами. В таком случае настройку или расширение модели данных следует проводить физически, на стадии проектирования.

Как и обычные БД семейства SIS, модель Seabed покрывает все потребности отрасли, уменьшая тем самым необходимость переноса данных. Открытый интерфейс SQL позволяет «на лету» преобразовывать модули, не меняя их код. Исходные данные хранятся в Больших Двоичных Объектах (Binary Large Objects, BLOBS), допускающих как последовательные, так и произвольные доступ и индексацию.

Объем данных, которые можно хранить, включая даже изображения, погрешности и гистограммы, которые обычно не хранят в реляционных БД, теоретически не ограничен. Для работы с пространственными привязками данных и с проекциями в систему включены функции ArcSDE, разработанные компанией ESRI.

Таким образом, и в приложениях и при работе с данными можно использовать средства геоинформационных систем.

Модель Seabed предоставляет возможности, которых пять лет назад еще не было. Это полнофункциональное средство хранения с улучшенными характеристиками безопасности, гибкости, расширяемости, масштабируемости, модульности и синхронизации и с дополнительными средствами улучшит работу отрасли, где используются большие потоки данных.

Управление данными в программах MMS

В ноябре Управление морских полезных ископаемых (Offshore Minerals Management, OMM) Службы полезных ископаемых (Minerals Management Service, MMS) МВД США приняло модель Seabed за основу проекта освоения внешнего континентального шельфа (Outer Continental Shelf Connect Project).

Этот проект является многоэтапным долгосрочным планом преобразования OMM в рамках концепции «электронного правительства» (e-Government). Проект нацелен на улучшение основных процессов программ OMM, включая замену действующих схем управления информацией современными программными продуктами. Предполагается, что таким образом удастся удовлетворить нужды пользователей и их групп, таких как органы управления (федеральные, уровня штата и местные), частный бизнес, научное сообщество, международные организации и общественность.