

ОБЪЕДИНЕННЫЙ  
ИНСТИТУТ  
ЯДЕРНЫХ  
ИССЛЕДОВАНИЙ

Дубна

95-50

P10-95-50

К.Ф.Окраинец

АВТОМАТИЧЕСКИЙ ЭКСПОРТ  
РЕЛЯЦИОННЫХ БД В WWW

Направлено на Международную конференцию «WWW-95»,  
Германия, 1995

1995

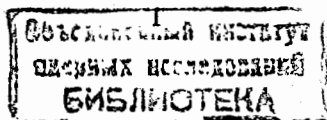
# 1 Введение

Все возрастающая популярность и богатые презентационные возможности распределенной гипертекстовой системы WWW [1] приводят к настоятельной необходимости предоставить пользователю то огромное количество информации, которое накоплено и хранится в разнообразнейших базах данных (БД). В данной работе описан один из возможных вариантов решения для больших распределенных реляционных БД. Результат доступен как <http://weirdb.jinr.dubna.su/wora/wora.html>  
<http://dozer.us.oracle.com:8080/sdk10/wora/>

# 2 Конфигурация

Дальнейшее изложение относится конкретно к следующей конфигурации:

- платформа сервера БД - UNIX System V;
- БД - Oracle;
- язык запросов - Oracle-диалект SQL;
- сетевая среда - TCP/IP Интернет;
- клиент БД - произвольное Oracle-приложение на произвольной платформе;
- W3-клиент - произвольный WWW-клиент (browser), способный интерпретировать формы (fill-out forms).



Однако, без потери общности, можно заменить операционную платформу сервера БД (Oracle имеет единый интерфейс для всех ОС и реализован практически на всех платформах), самое БД (другие лидеры в этой области имеют очень похожую архитектуру, например, Informix, Sybase) и сетевую среду (функциональные характеристики не зависят от способа соединения с БД).

### 3 Основные понятия

#### 3.1 Просто БД

Обычный способ извлечения информации из базы данных состоит в следующем: клиент соединяется с БД, используя либо коммерческое приложение, которое предоставляет фирма-производитель базы, либо приложение, которое разработано с помощью библиотек, опять-таки коммерческого характера. После соединения клиент выполняет некоторое количество запросов и отсоединяется от БД.

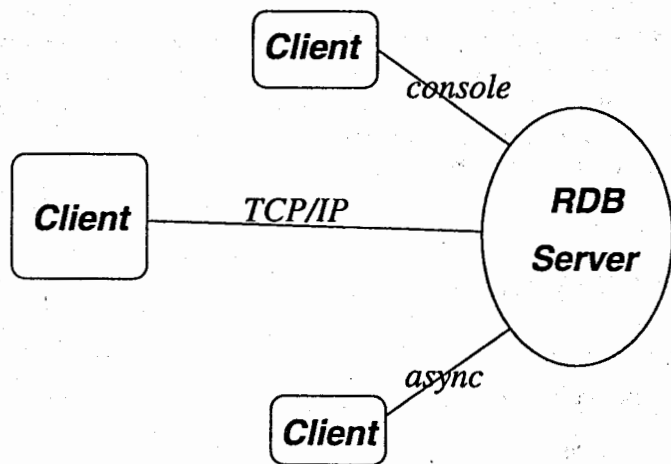


Рис. 1. Обычное соединение с БД

#### 3.2 Просто WWW

Где-то в сети существует HTTP-сервер (WWW), который имеет доступ к дереву документов, как правило, хранящихся в локальной для него файловой системе. Среди документов могут быть исполняемые программы (CGI-script), которые при вызове генерируют гипертекстовые файлы, выдавая их на свой стандартный выход (а сервер ретранслирует их клиенту). W3-клиент соединяется с HTTP-сервером и запрашивает некий документ. HTTP-сервер извлекает документ и передает его W3-клиенту, после чего последний отсоединяется.

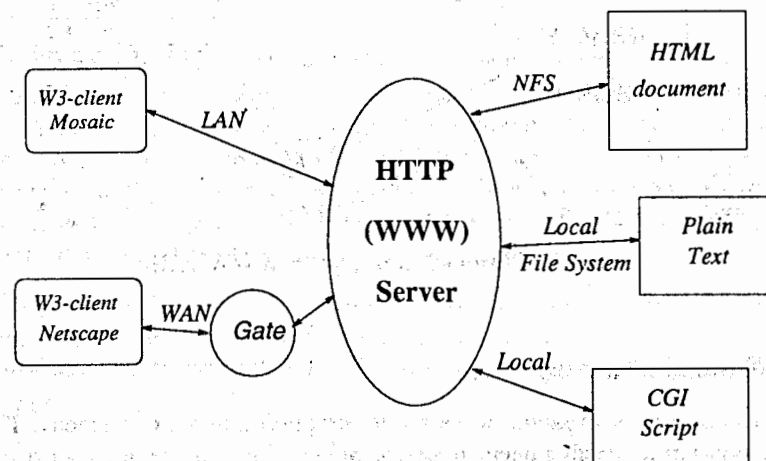


Рис. 2. Обычное соединение с HTTP-сервером

#### 3.3 WWW + БД

Между HTTP-сервером и БД появляются заполняемая форма (fill-out form) и исполняемая программа(ы) (CGI-script), к которым и обращается клиент в своих запросах. Форма представляет собой элемент языка HTML [2], который дает пользователю возможность задавать некоторые параметры (например, условия поиска в БД). При этом форма должна содержать ссылку на исполняемую программу, которой средствами операционной среды и будут передаваться заданные пользователем ограничения. Программа формирует из данных ей па-

параметров запрос, тем или иным способом соединяется с БД, передает ей запрос, принимает результаты, преобразует их в HTML-документ и выдает результат на стандартный выход.

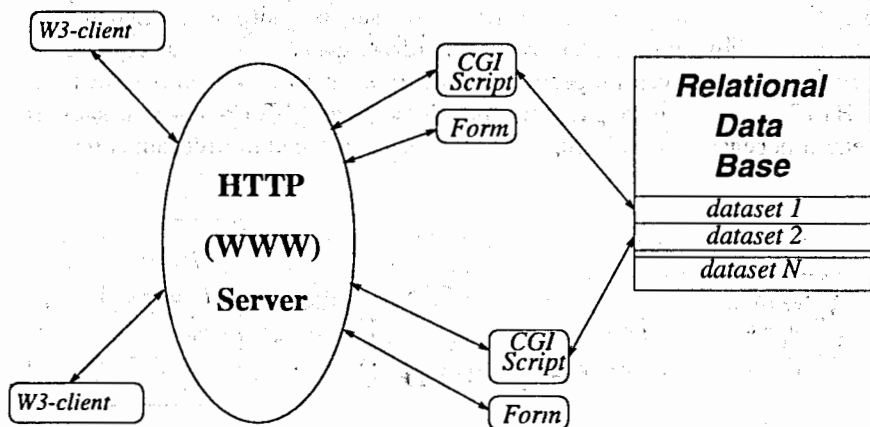


Рис. 3. Обычный интерфейс WWW с БД

### 3.4 Способы интеграции

Общая схема интеграции остается неизменной при всех способах. К сожалению, такая организация очень неэффективна. Еще одна неприятность состоит в том, что HTTP - это протокол, не удерживающий состояния (stateless), то есть нет способа извлечь несколько документов за одно соединение, нет способа заставить сервер помнить о состоянии клиента. Наиболее популярный [3] способ предоставить содержимое базы пользователям WWW состоит в следующем:

- описать форму, в которой пользователь будет задавать параметры;
- написать программу, которая будет анализировать параметры и выполнять запрос к БД;
- разместить форму и программу в дереве документов HTTP-сервера.

Эта последовательность действий должна быть повторена каждый раз, когда нужно экспортировать, сделать доступными через WWW новые данные или когда структура уже экспортированных данных меняется. Помимо очевидной головной боли для человека, который должен все это проделывать, такой

подход ограничивает возможности пользователя несколькими опциями, которые задал создатель формы. Таким образом, вся мощь SQL (Structured Query Language) остается невостребованной.

## 4 Архитектурное решение и реализация

### 4.1 Отличия, достоинства и недостатки

Итак, для того, чтобы преодолеть недостатки, свойственные традиционным способам интеграции WWW с БД, был разработан и реализован нижеописанный алгоритм.

Предлагаемое решение позволяет намного лучше использовать факт общения с реляционной БД. Помимо того, упрощается администрирование экспорта данных. Точнее, администрирование просто исчезает. Все, что нужно администратору БД для экспорта данных в WWW, это набрать строку, дарующую привилегию выборки из данной (виртуальной) таблицы псевдопользователю. При этом администратор БД может абсолютно ничего не знать о WWW.

Недостатками реализации являются:

- Единый интерфейс ко всем данным. Обычно это ставится в заслугу, но хотелось бы иметь возможность выборочно улучшать зрелищность форм и извлекаемых данных. Впрочем, это недостаток именно реализации, но не подхода;
- Отсутствие возможности вводить "сырые" (raw) SQL-операторы. Это легко можно сделать, но при этом возникнет угроза несанкционированного доступа к данным;
- Ограничения на вводимые условия выборки. Хотя это и не полный доступ к возможностям БД, но уже шаг вперед.

### 4.2 Основная идея

Как известно, у многих БД в той или иной форме существует внутренняя структура данных, в которой содержатся сведения о таблицах, пользователях, зачастую об операторах, типах данных и т.п. Назовем ее словарем (dictionary), тем более что она почти всегда так и называется.

Программа-скрипт (далее WORA, Wwww-ORacle), работая со словарем, должна извлекать из него все, что потребуется. Кроме того, она сама должна генерировать формы; заполнив которые, пользователь будет обращаться к ней

же. То есть одна программа выполняет все функции взаимодействия с БД и, косвенно, с пользователем, причем для любых данных.

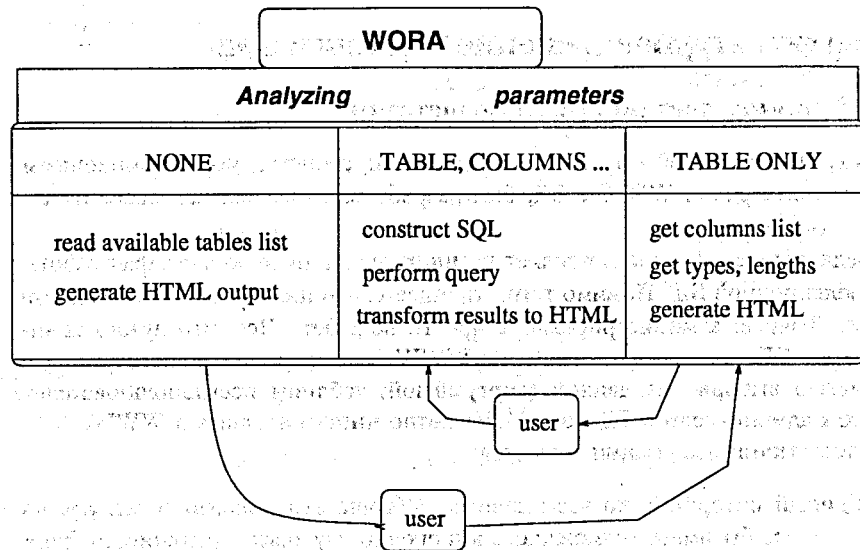


Рис. 4. Алгоритм WORA

При запуске программа соединяется с БД, используя идентификатор специально заведенного псевдопользователя. Далее поведение WORA таково:

1. Нет параметров - WORA извлекает из словаря список доступных ей (а следовательно, любому WWW-клиенту) таблиц и генерирует гипертекст с ссылками на эти таблицы.
2. Задано только имя таблицы - WORA извлекает из словаря список имен элементов заданной таблицы, их типов, размеров и функций SQL, применимых к каждому элементу. Имея все эти данные, WORA генерирует свою базовую форму, в которой пользователь может задать список извлекаемых полей, функций, применяемых к этим полям, и критерий поиска. Это случается, когда пользователь активировал какую-то из ссылок, упомянутых в предыдущем пункте.
3. Заданы имя таблицы и имена полей - WORA конструирует SQL-запрос и, используя Dynamic SQL Method 4, посылает его БД. Ответ базы транслируется W3-клиенту в виде HTML-документа. Это значит, что активи-

рована базовая форма, то есть пользователь выполняет непосредственно запрос данных.

Следует отметить, что обработка ошибок выполняется БД, и код ошибки с комментарием возвращается W3-клиенту. Это не всегда удобно, поскольку пользователь не обязательно понимает специфику БД. Для того, чтобы уменьшить количество возникающих недоразумений, вокруг WORA создан комплекс гипертекстовых документов, поясняющих особенности Oracle, SQL, тонкости с локализацией и некоторые возможные ошибки.

Рис. 5. Пример базовой формы

## WORA (Www-ORacle gate) at JINR

sunct0.jinr.dubna.su connected to DB via WORA

### Raw SQL statement

```
select (TABLE_NAME), (EXPLANATION) from HTTPD.DESCRPTIONS
```

### Select phase

- OCR.SERVICES, This is table in which port numbers, both in character and in numerical forms, IP protocol and network service name are stored.,
- OCR.DUBNA\_PHONES, Local phone numbers database. Field names are self-explanatory. Contents is in KOI-8 coding. If your comp is not russified – see WORA docs on how to help it.,
- HTTPD.DESCRPTIONS, About itself – containing short explanation what table is what about,

3/0 records/Kb selected  
Goodbye

Рис. 6. Пример вывода

## Список литературы

- [1] T. Berners-Lee, R. Gailliau et al.: *World Wide Web: the Information Universe*  
Electronic Networking 2(1)52-58, 1992.
- [2] T. Berners-Lee, D. Connolly: *HyperText Markup Language (HTML)*,  
URL=<http://info.cern.ch/hypertext/WWW/MarkUp/MarkUp.html>, 1993.
- [3] J. Ng: *GSQl*,  
<http://www.ncsa.uiuc.edu/SDG/People/jason/pub/gsql/starthere.html>

Рукопись поступа в издательский отдел  
9 февраля 1995 года.

Окраинец К.Ф.

P10-95-50

Автоматический экспорт реляционных БД в WWW

Все возрастающая популярность и богатые презентационные возможности распределенной гипертекстовой системы WWW приводят к настоятельной необходимости предоставить пользователю то огромное количество информации, которое накоплено и хранится в разнообразнейших базах данных (БД). В данной работе описан один из возможных вариантов для больших распределенных реляционных БД. Результат доступен как <http://weirdb.jinr.dubna.su/wora/wora.html>.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Препринт Объединенного института ядерных исследований. Дубна, 1995

Перевод автора

Осраинетс С.Ф.

P10-95-50

Auto-Exporting Relational Databases to WWW

Staggering popularity and advanced presentation capabilities of the World Wide Web make it necessary to give W3 users easy access to data stored in relational databases. Article describes an approach for automatic content-providing with database-sourced pages and its implementation for Oracle databases on genetic Unix platform. Results are accessible as <http://weirdb.jinr.dubna.su/wora/wora.html>.

The investigation has been performed at the Laboratory of Computing Techniques and Automation, JINR.

Preprint of the Joint Institute for Nuclear Research. Dubna, 1995