

РАСПРЕДЕЛЕННАЯ СИСТЕМА ОБРАБОТКИ ДАННЫХ В СИСТЕМАХ МОНИТОРИНГА ОКРУЖАЮЩЕЙ СРЕДЫ

А.Е.Нечесин, П.Л.Шугрин

Севастопольский национальный
технический университет
г.Севастополь, Студгородок
E-mail: tk@sevgtu.sebastopol.ua

Предложена распределенная система обработки данных, поступающих в региональные центры экологического мониторинга Украины.

В современных системах контроля окружающей среды, осуществляющих региональный мониторинг и работающих в реальном масштабе времени, генерируется множество временных рядов большой размерности. Это приводит к определенным сложностям при формировании первичных записей в базах данных. Данные измерений, поступающие с регионально распределенных источников, как правило, обладают свойством избыточности, содержат случайные выбросы, а также имеют неоднородный характер. Поэтому возникает задача разработки распределенной системы обработки данных (РСОД), позволяющей эффективно обрабатывать поступающие потоки разнородных данных с целью выделения информации, определяющей принятие решения лицом, управляющим процессами сбора и обработки данных и представляющим результаты мониторинга на более высокий уровень рассмотрения.

В настоящее время в областных центрах Украины организованы региональные центры экологического мониторинга [1]. Эти центры осуществляют систематическое наблюдение, сбор, обработку и хранение данных о состоянии окружающей среды в регионе. При этом общий информационный поток включает множество данных по источникам загрязнения, видам загрязняющих веществ, картографические данные, геогра-

фию загрязнения, данные радиозондовых наблюдений и наземных метеостанций.

Для более эффективной организации обработки и анализа данных, накапливаемых в региональных центрах, предлагается использовать информационную технологию распределенных баз данных (РБД). Применение РБД позволяет разместить часто используемые данные ближе к клиенту, что способствует минимизации сетевого трафика; расположить часто меняющиеся данные в одном месте, обеспечивая минимальные затраты по синхронизации копий данных; увеличить общую надежность системы к единичным отказам серверов (горячее резервирование); понизить стоимость системы за счет использования группы небольших серверов вместо одного мощного центрального [2, 3].

Исходя из территориально-распределенного характера размещения центров сбора данных, структуру распределенной системы обработки данных можно представить в виде набора узлов, связанных коммуникационной сетью, где каждый узел является полноценной СУБД. Структура РСОД приведена на рис. 1.

На этой схеме изображены узлы РБД, которые соответствуют региональным центрам со своими локальными базами данных и группами пользователей. Специальный компонент локальной СУБД узла предоставляет функциональные возможности для взаимодействия с другими СУБД системы. Каждый из пользователей системы может получать необходимую ему информацию как из БД только «своего» узла, так и из остальных БД, входящих в систему.

Схема информационных потоков в отдельном узле системы, выполняющей прием, хранение и обработку больших массивов разнородных данных на основе предварительной маршрутизации и буферизации, приведена на рис. 2.

Первичные данные от источников могут поступать различными способами, на-

пример в режиме реального времени через специальные каналы связи или в виде файла с накопленными за некоторый период данными на центральный ftp-сервер через глобальную сеть Интернет.

Процессами приема и распределения данных управляет коммуникационная система приема. Принятые данные класси-

фируются по формату и проходят два этапа первичной обработки, предусматривающие промежуточное хранение данных в буферных базах данных формата Dbase III. Алгоритмы первичной обработки данных реализованы в виде программных модулей на языках высокого уровня, использующих



Рисунок 1 – Структурная схема распределенной системы обработки данных

процессор баз данных CodeBase для доступа к буферным базам данных.

Первичные данные, прошедшие необходимую подготовку, поступают в базы оперативных данных, представляющие собой специализированные предметно-ориентированные часто обновляемые БД. Область их использования заключается в формировании отчетов по текущим данным, поддержки оперативных решений, а также консолидации данных из нескольких источников. Процессами загрузки данных в

базы оперативных данных управляет подсистема загрузки.

Основным местом хранения поступающих данных является хранилище данных. Это комплекс СУБД, на который возлагаются функции хранения больших объемов данных и обработки запросов на множестве этих данных. В связи с этим хранилище данных имеет специальную архитектуру, основанную на использовании СУБД MSSQL 2000. В рассматриваемом случае хранилище данных вместе с информационно-поисковой системой составляет основу

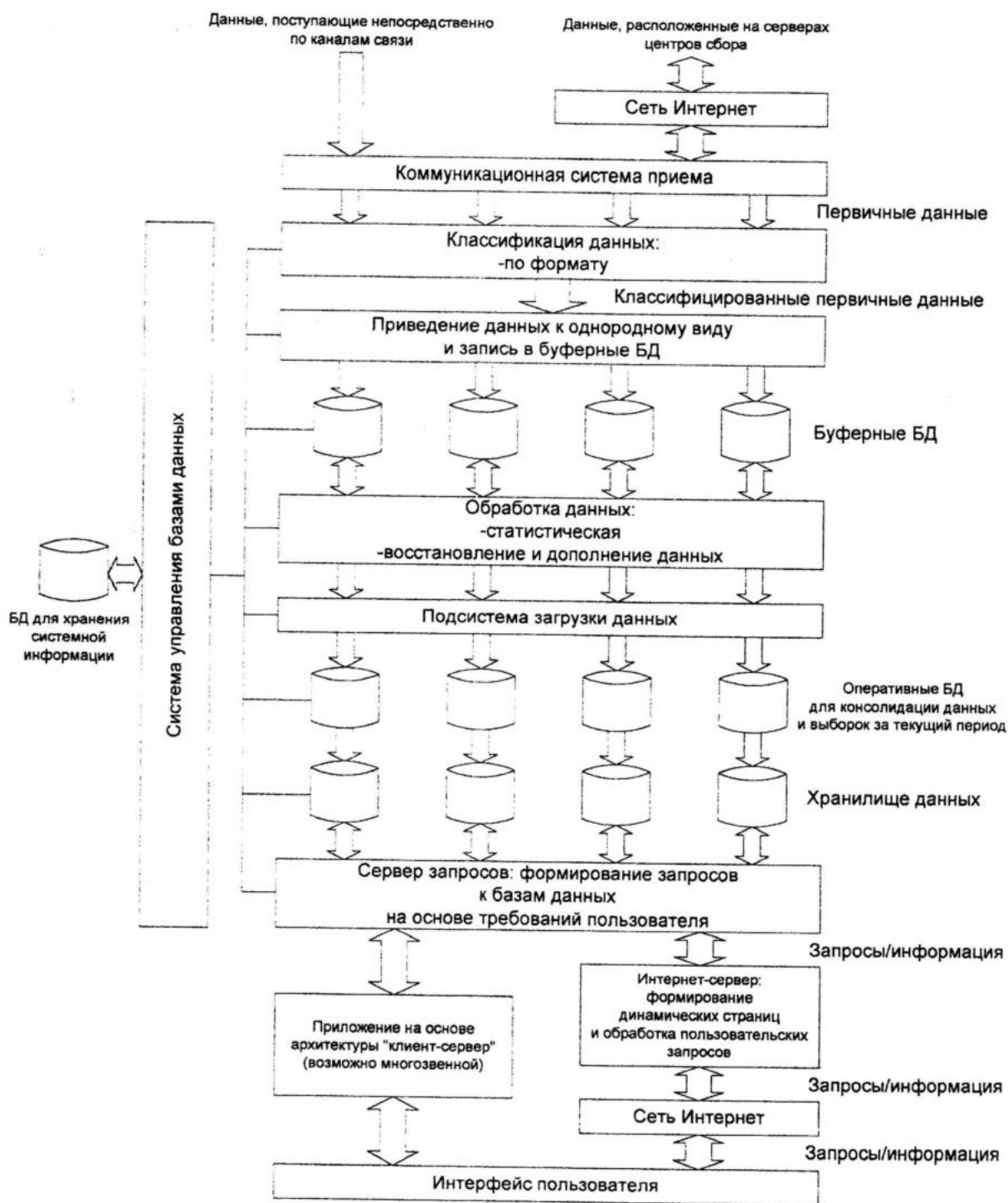


Рисунок 2 - Схема информационных потоков

системы принятия решений. При этом лицо, ответственное за принятие решения, имеет возможность получения информации за заданный интервал времени двумя способами - в виде некоторых обобщенных показателей или прогноза, а также в виде таблицы необходимых параметров. В первом случае выдача информации происходит непосредственно через web-интерфейс браузера с учетом некоторой задержки, необходимой для выполнения запроса к базе данных. Во втором случае файл с данными в формате, выбираемом самим пользователем, может быть прислан на электронный почтовый адрес пользователя, либо помещен на специальный ftp-сервер с возможностью загрузки на компьютер пользователя; при этом на электронный почтовый адрес пользователя высылается уведомление о возможности загрузки с указанием необходимых сведений.

Таким образом, РСОД позволяет достаточно эффективно решать вопросы сбора, хранения и обработки разнородных данных измерений, необходимых для организации экологического мониторинга в регионах Украины. Использование всемирной компьютерной сети Интернет для организации одного из вариантов пользовательского интерфейса имеет целый ряд преимуществ. В частности, облегчается доступ пользователей к информации (им достаточно иметь персональный компьютер с выходом в Интернет и программу-браузер), снимается проблема обновления клиентского программного обеспечения, т.к. при необходи-

мости внесения изменений в алгоритмы выдачи информации (например, добавление новых отчетов, графиков, диаграмм) меняются лишь программные модули на стороне сервера, которые расположены в региональных центрах экологического мониторинга.

Разработанная РСОД является открытой для расширения функциональных возможностей за счет создания проблемно-ориентированных приложений. Такие приложения позволяют специалистам различных областей использовать всю накопленную информацию в требуемых разрезах, за счет чего значительно расширится область применения системы.

ЛИТЕРАТУРА

1. Ульшин В.А., Акименко В.В., Тихонюк П.С., Докашенко А.И., Власов Ю.Н. Информационные системы анализа и прогнозирования загрязнения атмосферы стационарными источниками выбросов // Эко-технологии и ресурсосбережение. – 2000. – № 2. – С. 52-59.
2. Козленко Л., Проектирование информационных систем
(<http://www.interface.ru/ca/Proekt.htm>)
3. Дейт, К., Дж. Введение в системы баз данных: Пер. с англ. – М.: Издательский дом «Вильямс», 2001. – 1072 с.