

УДК 681.324

В.И. Шевченко, ассистент*Севастопольский национальный технический университет**ул. Университетская 33, г. Севастополь, Украина 99053**E-mail: kvf@sevgtu.sebastopol.ua***МОДЕЛЬ ОРГАНИЗАЦИИ ЭФФЕКТИВНОГО РАЗМЕЩЕНИЯ ТАБЛИЦ ДАННЫХ В КОРПОРАТИВНЫХ ИНФОРМАЦИОННЫХ СИСТЕМАХ**

Рассмотрена модель организации эффективного размещения таблиц данных в вычислительной сети корпорации, с учетом времени обслуживания запросов. Предложена программная реализация модели в среде имитационного моделирования AnyLogic.

Современные экономические условия требуют формирования новых подходов к процессам управления предприятиями. Одним из таких подходов является реинжиниринг бизнес-процессов на основе современных информационных технологий. Он ориентирован на реализацию принципов сквозного управления цепочками операций как на уровне подразделений предприятия, так и на межорганизационном уровне, позволяя встраивать бизнес-процессы поставщиков и потребителей продукции в совместно управляемые цепочки поставок [1, 2]. Реализация принципов реинжиниринга бизнес-процессов, как правило, влечет за собой изменения в архитектуре существующей корпоративной информационной системы предприятия (КИС). В этой связи актуальной является задача разработки моделей, позволяющих выбрать наиболее эффективную структуру взаимосвязи отдельных информационных систем в информационной сети предприятия.

Целью статьи является разработка модели, позволяющей ускорить обработку информационных потоков, связывающих участников деловых процессов и улучшить синхронизацию одновременно выполняемых бизнес-процессов за счет организации эффективного размещения таблиц баз данных КИС в локальных узлах корпоративной вычислительной сети. Обобщенно структура корпоративной сети предприятия может быть представлена схемой (рисунок 1), где U_i – узел обработки данных, ($i = 1, N$), U_{main} – центральный узел сети, соответствующий головному офису предприятия, U_{i_j} – локальный узел обработки данных, ($j = 1, m_i$).

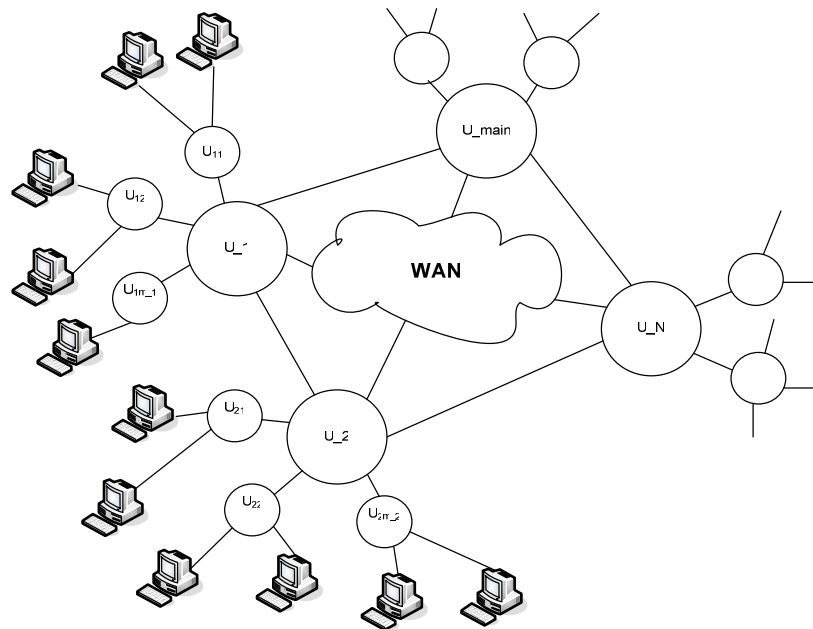


Рисунок 1 – Структура корпоративной сети предприятия

Для территориально распределенных организаций на логическом уровне информационный обмен в КИС реализуется как поток запросов к распределенным базам данных предприятия (РБД) и поток данных выборки информации по запросам. В этом случае эффективное управление в КИС поддерживается с помощью системы управления распределенной базой данных (УРБД), использующей механизмы фрагментации, тиражирования и репликации данных [3]. В случае функциональной распределенности предприятия, либо на уровне локальных узлов обработки данных, информационный

обмен в КИС может быть организован как обмен данными между обособленными информационными системами (ИС) одинаковой логической структуры. При этом с целью минимизации операций ввода идентичных по предприятию данных в соответствующие таблицы базы данных для каждой информационной системы, часть таблиц в информационных системах может быть обобщена и физически размещена на каком-либо из вычислительных узлов КИС и к ней организован общий доступ (рисунки 2, 3). Число и тематика таблиц баз данных общего доступа определяются на уровне реинжиниринга логической структуры КИС, как правило, это таблицы электронных справочников, содержащие информацию о качественных характеристиках номенклатуры, о клиентах бизнес-процессов и еще ряд вспомогательной информации. Частичное объединение таблиц данных ИС имеет как свои преимущества, так и недостатки. С одной стороны, сокращается время ввода данных общего пользования, с другой – увеличивается время обработки запроса к удаленной части информационной системы.

Предлагаемая модель организации размещения таблиц баз данных в КИС позволяет на этапах проектирования и реинжиниринга логической структуры информационной системы уровня предприятия выбрать наиболее эффективный по времени обработки запросов вариант размещения таблиц баз данных в локальных узлах вычислительной системы.

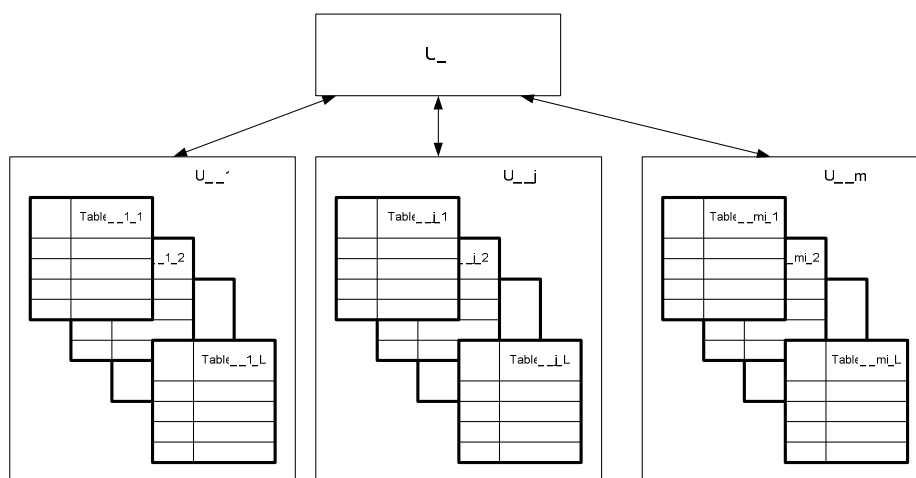


Рисунок 2 – Схема размещения таблиц данных до реорганизации

Для описания входных параметров модели введены следующие обозначения:

$i = \overline{1, n}$ – число информационных систем в КИС;

$j = \overline{1, m_i}$ – число локальных вычислительных узлов с обособленными ИС, участвующими в процессе обмена информацией;

$l = \overline{1, L}$ – число таблиц в базе данных ИС (предполагается, что каждый узел содержит ИС одинаковой логической структуры);

n_l – условная предельная мощность l -ой таблицы ИС (объем памяти, резервируемый под одну запись в l -ой таблицы ИС), байт;

$z = \overline{1, Z}$ – число типов запросов к таблицам ИС на выборку информации;

$$d_{1_zl} = \begin{cases} 1, & \text{если } z\text{-ый запрос использует } l\text{-ю таблицу данных} \\ 0, & \text{иначе} \end{cases} \quad (1)$$

d_{1_zl} – элемент матрицы соответствия «запрос-таблица» D_1 , определяемый по формуле (1)

$k = \overline{1, K}$ – число типов запросов на обновление таблиц ИБ, $K \leq Z$;

$$d_{2_lk} = \begin{cases} 1, & \text{если } k\text{-ый запрос использует } l\text{-ю таблицу данных} \\ 0, & \text{иначе} \end{cases} \quad (2)$$

d_{2_lk} – элемент матрицы соответствия «запрос-таблица» D_2 , определяемый по формуле (2)

λ_{1_z} – интенсивность запросов z -го типа на выборку информации;

λ_{2_k} – интенсивность запросов k -го типа на обновление данных в таблице;

μ_{1_z} – интенсивность обработки запросов z -го типа на выборку информации;

μ_{2_k} – интенсивность обработки запросов k -го типа на обновление данных единичной мощности (на 1 байт информации);

μ – интенсивность передачи данных единичной мощности. (В целях упрощения модели предполагается, что общие таблицы баз данных размещаются на независимом узле вычислительной системы и скорость доступа запросов от удаленных ИС для всех вычислительных узлов, на которых размещены ИС одинакова);

T_z, T_k – предельные времена реакции системы на запросы пользователя z -го и k -го типов к ИС.

Необходимо определить множество таблиц S , к которым будет организован общий доступ в ИС, при условии, что предельные времена обработки запросов не превысят заданные.

В общем случае, время обработки запроса на выборку данных в не связанных ИС: $t_{1_z} = 1/\mu_{1_z}$. Предположим, что операторы различных филиалов предприятия имеют равный уровень квалификации, при этом время обработки запроса на обновление ИС с учетом числа ИС в системе составит: $t_{2_k} = m/\mu_{2_k}$.

Для организации связи ИС посредством общих таблиц данных необходимо определить квадратную матрицу соответствия D_3 следующего вида:

$$d_{3_ls} = \begin{cases} 1, & \text{если } l\text{-я таблица является элементом множества } S \\ 0, & \text{иначе.} \end{cases} \quad (3)$$

В этом случае время обработки запроса на выборку данных определяется следующим образом:

$$t_{1_z} = \frac{1}{\mu_{1_z}} + \sum_{l=1}^L \sum_{s=1}^S \frac{d_{1_lz} d_{3_ls}}{\mu} \quad (4)$$

Время обработки запроса на обновление равно:

$$t_{2_k} = 1/\mu_{2_k}. \quad (5)$$

На этапе предварительного анализа группой экспертов может быть выделено множество таблиц данных S^t , к которым возможно организовать общий доступ. В ходе имитационных экспериментов над моделью получено множество $S \subseteq S^t$, для которого выполняются следующие условия:

$$\begin{cases} \frac{1}{\mu_{1_z}} + \sum_{l=1}^L \sum_{s=1}^S \frac{d_{1_lz} d_{3_ls}}{\mu} \leq T_z, \\ \frac{1}{\mu_{2_k}} \leq T_k. \end{cases} \quad (6)$$

Схема размещения таблиц ИС, после реорганизации изображена на рисунке 3.

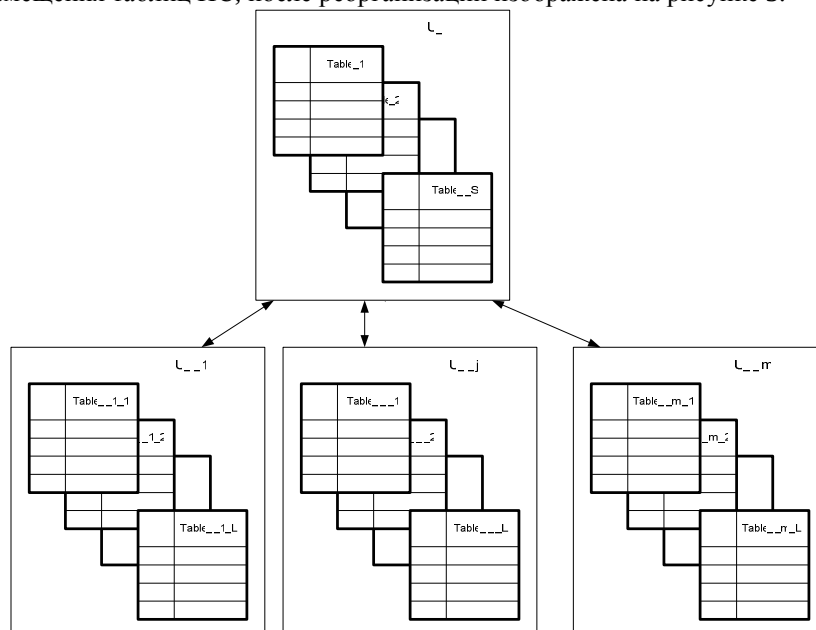


Рисунок 3 – Схема размещения таблиц данных после реорганизации

Рассмотрим работу модели на следующем примере: в системе предприятия функционируют три обособленных информационных системы, имеющих одинаковую логическую структуру и реализованных на базе компонента «Оперативный учет» системы «1С: Предприятие». Фрагмент структуры метаданных и файла словаря базы данных изображен на рисунке 4. В ходе предварительного анализа информационных систем выявлена возможность объединения трех из 28 таблиц ИС. Это элементы справочников: SC12 (Клиенты), SC13 (Товары) и SC18 (Районы). Входные данные для модели приведены в таблице 1.

Для разработки имитационной использована среда AnyLogic, стандартные компоненты библиотеки Enterprise Library. Данная среда представляет широкие возможности для моделирования и высокоуровневый интерфейс для быстрого создания дискретно-событийных моделей с помощью блочных схем по технологии drag-and-drop [4]. Модель ИС предприятия с интегрированными таблицами данных, реализованная в среде AnyLogic изображена на рисунке 5.

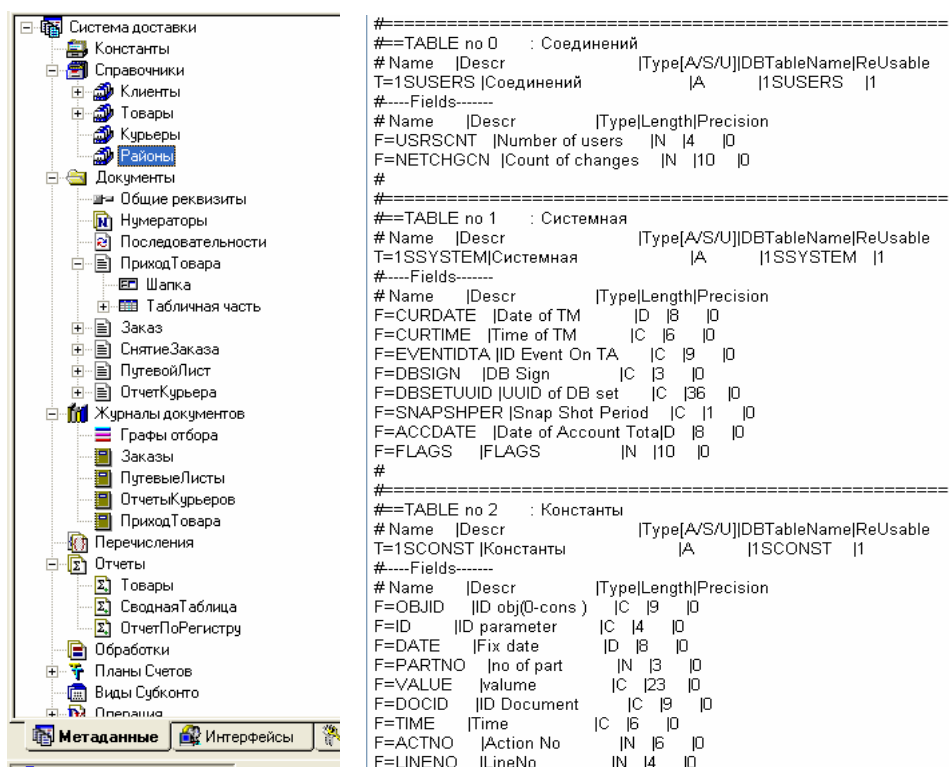


Рисунок 4 – Фрагмент структуры данных информационной системы

Таблица 1 – Входные данные для моделирования

Структурные характеристики КИС		Значения характеристик
Число локальных вычислительных узлов, участвующих в процессе обмена информацией		3
Число таблиц в базе данных ИС		28
Число таблиц, подлежащих объединению		3
Состав таблиц, подлежащих объединению:		SC12, SC13, SC18
Предельное время реакции ИС на запросы об обновлении информации		15 мин
Предельное время реакции ИС на запросы на выдачу информации		5 мин
Параметры моделирования	Среднее значение	Распределение
Интервал времени между появлениями запросов на обновление	20 мин	Экспоненциальное
Интервал времени между появлениями запросов на выборку данных	10 мин	Экспоненциальное
Время обработки запроса на обновление информации единичной мощности	3 с	Равномерное в интервале 2...4
Время передачи данных единичной мощности	20 с	Равномерное в интервале 15...25

Для реализации процесса генерации запросов к ИС в имитационной модели используется три блока *source* (APM1, APM2, APM3), связь в таблицах данных имитируют девять блоков *selectOutput*. Для моделирования общих таблиц доступа использованы 3 блока *queue* и 3 блока *delay* (SC12, SC13, SC18). Для упрощения модели, обращение к оставшимся 25 таблицам для каждой из трех ИС реализовано тремя блоками *delay* (Table_APM_1, Table_APM_2, Table_APM_3). Идентификация поступающего в систему запроса, как запроса на обновление и имитация запроса на обновление данных реализованы шестью парами блоков *selectOutput* и *delay*. Для сбора статистики об обработанных запросах в модели используются 3 блока *sink* (sink_APM1, sink_APM2, sink_APM3).

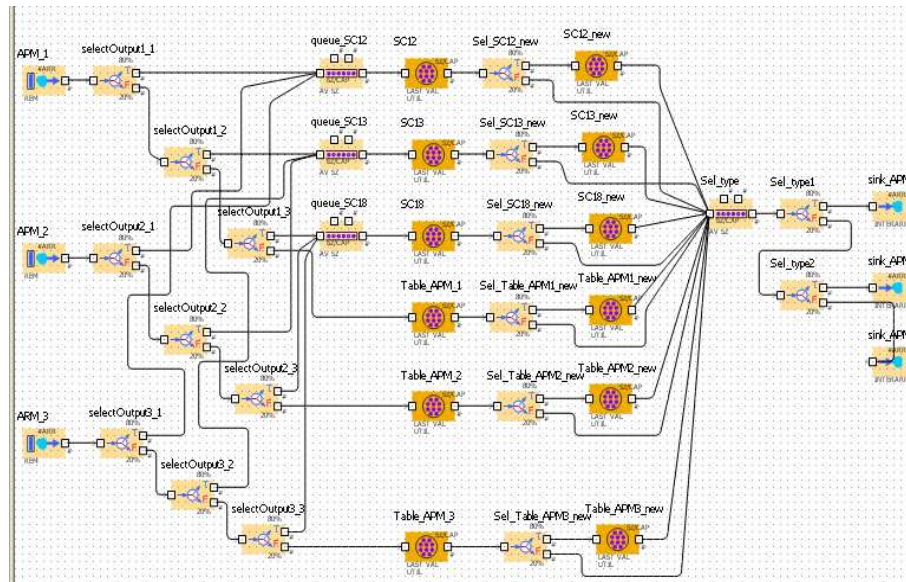


Рисунок 5 – Схема модели ИС предприятия в среде AnyLogic

Рассмотренная модель позволяет выбрать наиболее эффективный вариант размещения таблиц данных в узлах КИС с учетом времени обработки запросов. По результатам исследования для рассмотренного примера можно сделать вывод о целесообразности выделения двух таблиц SC12 (Клиенты) и SC13 (Товары) для использования в общем доступе тремя ИС с идентичной структурой конфигураций.

В перспективе дальнейших исследований предусматривается разработка моделей эффективного распределения таблиц данных, с учетом физических характеристик вычислительных узлов КИС, а также создание на основе данной и последующих моделей Java-апплетов, интегрируемых в структуру исследуемых информационных систем, в качестве модуля поддержки принятия решения в процессе реинжиниринга ИС.

Библиографический список

1. Репин В.В. Процессный подход к управлению. Моделирование бизнес-процессов / В.В. Репин, В.Г. Елиферов. — М.: Стандарты и качество, 2004. — 408 с.
2. Тельнов Ю.Ф. Реинжиниринг бизнес-процессов. Компонентная методология / Ю.Ф. Тельнов. — М.: Финансы и статистика, 2004. — 320 с.
3. Цегелик Г.Г. Системы распределенных баз данных / Г.Г. Цегелик. — Львов: Свит, 1990. — 168 с.
3. Карпов Ю. Имитационное моделирование систем. Введение в моделирование с AnyLogic5 / Ю. Карпов. — СПб.: БХВ-Петербург, 2005. — 400 с.

Поступила в редакцию 4.07.2008 г.