

УДК 681.518:378

Ю.Н. Возовиков¹, В.Я. Воропаева², В.Ф. Шапо³

- 1) Одесский национальный политехнический университет, г. Одесса
кафедра системного программного обеспечения
- 2) Донецкий национальный технический университет, г. Донецк
кафедра автоматизации и телекоммуникаций
- 3) Одесская национальная морская академия, г. Одесса
кафедра теории автоматического управления и вычислительной техники
E-mail: yuri_email@mail.ru, voropayeva@meta.ua, stani@te.net.ua

ВЫБОР ТОПОЛОГИИ И ОПТИМИЗАЦИЯ ЗАГРУЗКИ КОМПЬЮТЕРНЫХ СЕТЕЙ СПЕЦИАЛЬНОГО НАЗНАЧЕНИЯ

Аннотация

Возовиков Ю.Н., Воропаева В.Я., Шапо В.Ф. Выбор топологии и оптимизация загрузки компьютерных сетей специального назначения. Предложен метод выбора топологической структуры компьютерной сети специального назначения и оптимизации ее загрузки. Учтены цена, живучесть, объемы сетевого трафика, возникающего при обращении к удаленным СУБД в рамках территориально распределенной информационной системы предприятия.

Ключевые слова: компьютерная сеть специального назначения, топология, сетевой трафик, материализованное представление.

Общая постановка проблемы. После распада СССР и обретения Украиной независимости возникла необходимость в создании таможенной службы (ТС) практически заново. К существовавшим таможенным постам добавился ряд дополнительных, поскольку Украина получила новые границы с Россией, Молдовой и Беларусью. Всё активнее используются и создаются проходящие по территории Украины транспортные коридоры. Поэтому возникает необходимость в создании новых таможенных подразделений, расширении и модернизации имеющихся, а также в создании единой информационной системы (ИС), базирующейся на территориально распределенной компьютерной сети (ТРКС), отвечающей современным требованиям к аппаратному и программному обеспечению, сетевому и коммуникационному оборудованию. Решение этой задачи затрудняет нехватка средств и существенные ограничения в применении компьютерных сетей общего пользования в связи с закрытостью передаваемой информации. Поэтому невозможно или затруднено применение в чистом виде технологии VPN (Virtual Private Network), физически построенных на тех же линиях связи, что и другие сети, но использующих шифрование и каждый раз заново устанавливаемые логические соединения. Нельзя или нежелательно также применять протоколы Ipsec, https и аналогичные, использующие шифрование данных в сетях общего пользования. По тем же причинам невозможно применение модели SaaS, в которой пользовательские данные хранятся на серверах провайдеров хост-услуг. Кроме того, зачастую используются морально устаревшие низкоскоростные линии (каналы) передачи данных. В итоге ТРКС, на которой базируется ИС ТС, обладает низкой живучестью и недостаточной пропускной способностью, большим временем реакции системы на запросы пользователей при обращении к удаленным базам данных (БД), завышенной стоимостью, поскольку строится эвристически, на базе накопленного опыта подобных решений, без использования научного подхода [1 – 3].

Комплексный метод построения ТРКС общего назначения был рассмотрен в [3]. Данный метод может быть задействован и при построении ТРКС и ИС ТС.

Решение задач и результаты исследований. Предлагается метод построения ИС ТС, позволяющий формализовать и автоматизировать процесс их создания и модернизации и заключающийся в последовательном выполнении взаимосвязанных этапов. На первом этапе разрабатывается информационная модель (ИМ) ТРКС ИС на основе анализа связей между входящими в нее локальными компьютерными сетями (ЛКС), анализа типов и объемов циркулирующего в ней трафика и временного распределения нагрузок на отдельные узлы ИС.

Для анализа информационных потоков при разработке ИС ТС целесообразно использовать структурный анализ. В настоящее время известно около 90 его разновидностей [4, 5]. При этом используются три группы средств: DFD (Data Flow Diagrams) – диаграммы потоков данных или SADT (Structured Analysis Design Technique) – диаграммы, иллюстрирующие функции системы; ERD (Entity-Relationship Diagrams) – диаграммы "сущность-связь", моделирующие отношения между данными; STD (State Transition Diagrams) – диаграммы переходов состояний, моделирующие поведение системы во времени.

При анализе систем обработки информации обычно используется методология DFD [4, 5]. Практически любые системы могут быть промоделированы при помощи DFD-ориентированных методов, поэтому для построения ИМ ИС ТС выберем методологию DFD как в наибольшей степени соответствующую поставленной задаче.

Исходя из анализа типов и расчета объемов трафика, строятся модифицированные DFD-диаграммы D и множество двудольных графов $Z_G = \{G_1, G_2, \dots, G_i, \dots, G_z\}, D \rightarrow Z_G$, характеризующих обмен данными между ЛКС через ТРКС. Узлы ЛКС генерируют трафик нескольких типов. Трафик i -й ЛКС, $i = 1, \dots, z$, описывается двудольным графом; L_1, \dots, L_z – ЛКС, входящие в ТРКС; n_i – число узлов ЛКС $_i$; U_{1i}, \dots, U_{ni} – n узлов i -ой ЛКС.

Полный трафик через ТРКС

$$W_{RS} = \sum_{c=1}^z \sum_{c=1}^z \sum_{b=1}^{n_i} \left(\sum_{i=1}^k W_{Rabic} + \sum_{j=1}^m W_{Sabjc} \right), \quad (1)$$

где k и m — количество типов трафика реального времени (ТРВ) и синтетического трафика (ТСТ) соответственно.

Пропускная способность канала связи ТРКС B_m выбирается на базе рассчитанной B_{ij} [3] с учетом (1): имеется q значений стандартных пропускных способностей каналов связи, отсортированных по возрастанию, $B = (B_1, B_2, \dots, B_m, \dots, B_q)$. Тогда $\forall i, \exists B_m, B_{m-1} < B_{ij} \ \& \ B_m > B_{ij}$.

При выборе каналов связи необходимо учесть следующие варианты.

1. Сети общего пользования (Интернет) и сети специального назначения, использующие кабельные системы.
2. Сети, базирующиеся на спутниковых каналах передачи данных (Инмарсат и аналогичные) или на протоколах мобильной связи (протоколы 3-го и 4-го поколений).
3. Беспроводные сети стандарта IEEE 802.16: IEEE 802.16a, IEEE 802.16-2004.

Подходы, перечисленные в п. 1-3, могут быть применены при условии надлежащего шифрования передаваемых данных, защиты каналов передачи данных от прослушивания и соответствия требованиям жесткого реального времени.

При построении ТРКС чрезвычайно важным является построение ее топологии с учетом требований к живучести и пропускной способности линий связи (ЛС), так как неоптимальное проектирование приводит к излишним материальным и временным затратам. Синтез топологической структуры ТРКС выполняется с учетом размещения ее узлов, особенности прокладки ЛС и накладываемых оборудованием ограничений, требований к живучести и стоимости ТРКС. Информация о размещении узлов может задаваться с помощью системы координат или указанием расстояний между всеми узлами, что позволяет учесть конкретные особенности прокладки ЛС (мосты, водоемы, городская черта). Минимизируемым критерием

при построении ТРКС является ее стоимость. Представим ТРКС в виде графа (узлы – в виде вершин, ЛС – в виде ребер). Эта задача – частный случай задачи построения экстремального дерева (ЭД). В качестве веса ребра графа используем стоимость ЛС, соединяющей его вершины. Часто возникает необходимость модернизировать имеющуюся сеть, когда некоторые узлы уже соединены. При добавлении ребра учтем максимальную связность узла, определяемую аппаратными (число портов маршрутизатора или шлюза) и финансовыми ограничениями. В связи с неблагоприятным географическим положением узла некоторые ЛС не могут быть проложены и исключаются из рассмотрения до этапа введения связей в граф. Такая ситуация требует пересмотра ограничений или рассмотрения каждой полученной подсети по отдельности и связывания их по радиоканалу или спутниковой связи.

На втором этапе проектирования разрабатывается топологическая структура ТРКС $S_T = f(C, E)$ с учетом заданного при постановке задачи проектирования уровня живучести и соотношения цена C / живучесть E .

Метод повышения живучести ТРКС содержит следующие шаги.

1. Вычислить число ЭД графа сети m

$$m = C_{n-1}^k = \frac{(n-1)!}{k!(n-1-k)!}, \tag{2}$$

где k – количество возможных разрывов линий связи; n – число вершин графа сети.

2. Сформировать список ограничений из k связей.

3. Построить ЭД графа сети без использования связей, попавших в список, полагая, что такие соединения не могут существовать.

4. Повторить m раз шаги 2 и 3, после чего из всех полученных ЭД графа сети наложением получить результирующий граф.



Рисунок 1 – Граф сети, соединяющей узлы ИС ТС (повышенная живучесть)

Предложенный метод позволяет реализовать решение ряда задач, возникающих при построении ТРКС. Построение топологии сети с заданной живучестью и минимальной стоимостью, расчет времени транзакции между ее узлами и выбор сетевого оборудования позволяют избежать излишних финансовых и временных затрат при построении сети (рис. 1), где черным цветом указаны ЛС ЭД, а зеленым – дополнительно построенные ЛС, позволяющие повысить живучесть ТРКС и быстродействие ИС.

В работах [5, 6] показана эффективность управления включением/выключением материализованных представлений (МП) при работе ИС, поскольку одним из важнейших типов трафика являются запросы к удаленным БД. Однако вопрос определения моментов включения/выключения МП остается открытым. Эффективность МП может быть значительно повышена, если учитывать периодичность появления соответствующих запросов. Для исследования реальной ИС будем исследовать ее журнал транзакций. В работе [7] показано получение множества групп запросов G из журнала транзакций, для которых целесообразно строить МП и определять периодичность их включения и выключения $G = \{G_1, \dots, G_i, \dots, G_n\}$.

Далее, на примере группы запросов G_i , найдем периодичность их появления. Затем найдем периодичность для каждой группы множества G . Ниже предложен метод определения периода включения МП, основанный на анализе длительности включения МП и интервала между моментами эффективного использования МП.

Вначале определим эффективность применения МП для группы запросов G_i в течение периода наблюдения T как отношение времени выполнения всех запросов без внедрения МП ко времени выполнения этих же запросов при использовании МП:

$$Eg_i = \frac{\sum_{l=1}^k t_{Gil}}{\sum_{l=1}^k t_{mpi} + \sum_{l=1}^{k_u} t_{uil} + \sum_{l=1}^k t_z + \sum_{l=1}^{k_{nz}} t_{nz}} - 1, \tag{3}$$

где t_{Gil} – время выполнения всех запросов группы G_i за время t_p без внедрения МП; t_{mpi} – время выполнения всех запросов, которые входят в G_i за время t_p , при использовании МП; k_u – число обновлений МП за время t_p (число операций Update); k – число запросов типа Select; k_{nz} – число запросов типа Select, не вошедших в МП; t_{nz} – время выполнения запросов типа Select, не вошедших в МП; t_{uil} – время обновления соответствующего МП; t_z – время анализа запросов; t_p – минимальный интервал времени при определении плотности запросов.

На рис. 2 приведены графики, иллюстрирующие распределение во времени функции эффективности применения МП и управляющих сигналов МП.

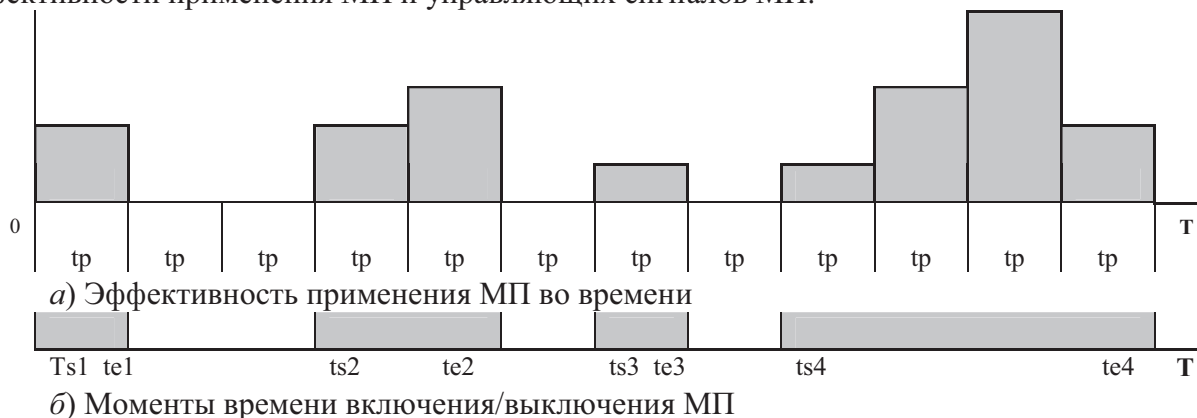


Рисунок 2 – Применение МП

Для оценки качества выбора периода и моментов включения МП введем следующие характеристики.

1. Максимально возможная эффективность применения МП – отношение времени выполнения всех запросов без внедрения МП ко времени выполнения этих же запросов с управлением вкл/выкл, в соответствии с рис. 1, б.

$$Eg_{\max i} = \frac{\sum_{l=1}^k t_{Gil}}{\sum_{l=1}^k t_{mpi} + \sum_{l=1}^{k_u} t_{uil} + \sum_{l=1}^k t_z} - 1. \quad (4)$$

2. Минимально возможная эффективность применения МП – эффективность применения МП при включении МП на весь период работы ИС.

Эффективность применения МП для группы запросов G_i при постоянно включенном МП получим минимально возможную эффективность применения МП:

$$Eg_{\min i} = \frac{\sum_{l=1}^k t_{Gil}}{\sum_{l=1}^k t_{mpi} + \sum_{l=1}^{k_u} t_{uil} + \sum_{l=1}^k t_z} - 1. \quad (5)$$

3. Средняя эффективность применения МП – рассчитываемая эффективность применения МП при средних расчетных величинах периода и длительности использования МП.

3.1. Средняя длительность включения МП:

$$tdl_{cp} = \sum (t_{ei} - t_{si}) / m, \quad (6)$$

где t_i – время использования каждой группы запросов G_i ($t_i = t_{ei} - t_{si}$); $T_{si} = \{t_{s1}, t_{si}, \dots, t_{sn}\}$ – моменты включения МП; $T_{ei} = \{t_{e1}, t_{ei}, \dots, t_{en}\}$ – моменты выключения МП; m – число расчетных периодов включения/ выключения МП.

3.2. Средний период включения МП:

$$tp_{cp} = \sum (t_{si-1} - t_{si}) / m, \quad (7)$$

$$Eg_{sr_i} = \frac{\sum_{l=1}^k t_{Gil}}{\sum_{l=1}^k t_{mpi} + \sum_{l=1}^{k_u} t_{uil} + \sum_{l=1}^k t_z + \sum_{l=1}^{k_{nz}} t_{nz}} - 1 \quad (8)$$

При $Eg_{\max i} \approx Eg_{\min i}$ или $Eg_{\max i} \approx Eg_{sr_i}$ можно рекомендовать использовать систему при постоянно включенном МП. При средних величинах ширины tdl_{cp} и периода tp_{cp} , рассчитывая Eg_{cp_i} , определим момент первого включения МП. При наибольшей величине Eg_{cp_i} (8) фиксируем момент включения МП. Варьируя величину ширины tdl_{cp} от min до max , определяем наибольшую величину Eg_{cp_i} (8) и оптимальную величину tdl_{cp} (5). Аналогично определяем оптимальную величину периода tp_{cp} (7). При новых величинах tdl , tp повторно определяем момент первого включения МП. Если расчетная величина $Eg_i > Eg_{cp_i}$ (8), то найденные моменты времени включения/выключения МП можно считать оптимальными. Поиск момента первого включения МП представляет собой итерационный процесс.

Выводы.

1. Предложен алгоритм выбора топологической структуры компьютерной сети специального назначения, позволяющий повысить ее живучесть.
2. Разработан метод расчета схемы управления материализованными представлениями, основанный на анализе эффективности использования МП с учетом периодичности работы ИС.
3. Обосновано использование схемы управления МП в различных информационных системах для уменьшения сетевого трафика, нагрузки на сетевое оборудование и стоимости ИС.

Литература

1. Клименко О. DataCom стремительно развивает всеукраинскую сеть передачи данных / О. Клименко // Компьютерное обозрение. – 2001. – №39. – С. 48 – 49.
2. Шапо В.Ф. Методы проектирования региональной вычислительной сети водного транспорта / В.Ф. Шапо // Вестник национального технического университета "ХПИ". – 2005. – Вып. 55. – С. 13 – 18.
3. Вендров А.М. CASE – технологии / А.М. Вендров. – М.: Финансы и статистика, 1998. – 175 с.
4. Калянов Г.Н. CASE: структурный системный анализ / Г.Н. Калянов. – М: Лори, 1996. – 246 с.
5. Шапо В.Ф. Оптимизация распределения запросов в компьютерных сетях специального назначения / В.Ф. Шапо, Ю.Н. Возовиков // Стратегия качества в промышленности и образовании: труды 6-й международной конференции. – Варна: технический университет. – 2010. – Т.2 (1). – С. 649 – 652.
6. Кунгурцев А.Б. Поиск закономерностей в распределении запросов для управления материализованными представлениями / А.Б. Кунгурцев, Ю.Н. Возовиков // Труды Одесского политехнического университета. – 2008. – 2.(30). – С. 135 – 140.

Надійшла до редакції:
16.03.2011

Рекомендовано до друку:
д-р техн. наук, проф. Скобцов Ю.О.

Abstract

Vozovikov Y.N., Voropaeva V.Y., Shapo V.F. Topology selection and loading optimization of special assignment computer networks. Method of special assignment network topology structure selection and loading optimization is proposed. Cost, vitality and network traffic volume of databases control systems in territorial distributed information system are registered.

Keywords: *special assignment computer network, topology, network traffic, materialized view*

Анотація

Возовиков Ю.Н., Воропаева В.Я., Шапо В.Ф. Вибір топології і оптимізація завантаження комп'ютерних мереж спеціального призначення. Запропоновано метод вибору топологічної структури комп'ютерної мережі спеціального призначення і оптимізації її завантаження. Враховано ціну, живучість, об'єми мережевого трафіку, що створюється під час звертання до віддалених СУБД в рамках територіально розосередженої інформаційної системи підприємства.

Ключові слова: *комп'ютерна мережа спеціального призначення, топологія, мережевий трафік, матеріалізоване уявлення.*

© Возовиков Ю.Н., Воропаева В.Я., Шапо В.Ф., 2011