

# ОПТИМАЛЬНОЕ ОПЕРАТИВНОЕ РЕЗЕРВИРОВАНИЕ ИНФОРМАЦИИ В СИСТЕМАХ ОБРАБОТКИ ДАННЫХ НА БАЗЕ ВЫЧИСЛИТЕЛЬНЫХ СЕТЕЙ

Е.А. Микрин, С.К. Сомов

Рассмотрены вопросы применения оперативного резервирования информации в распределенных системах обработки данных (РСОД), функционирующих в рамках вычислительных сетей, в целях повышения безопасности их работы. Проанализированы особенности применения трех основных стратегий оперативного резервирования в РСОД. Сформулированы задачи оптимального по разным критериям распределения оперативного резерва РСОД по узлам вычислительной сети и предложены методы их решения. Рассмотрены проблемы и задачи оптимизации резервирования массивов данных с учетом их взаимосвязи между собой и резервирование данных в распределенных базах данных. Выполнен анализ влияния трафика запросов на распределение резерва РСОД.

**Ключевые слова:** распределенные системы обработки данных, вычислительные сети, резервирование массивов данных, задачи оптимизации резервирования, стратегии резервирования.

## ВВЕДЕНИЕ

Распределенные системы обработки данных (РСОД) представляют собой сложные аппаратно-программные комплексы или системы, компоненты которых распределены на больших расстояниях друг от друга и объединены в единую систему различными каналами связи и специальным системным программным обеспечением. Они строятся, как правило, на основе глобальных, корпоративных или сетей специального назначения.

В процессе работы РСОД могут происходить инциденты, приводящие к появлению ошибочных данных, ошибочным результатам решения задач и обработки запросов системой, а в крайнем случае и к невозможности нормального функционирования РСОД, т. е. к потере системой работоспособности, ее отказу. Разрушение программного и информационного обеспечения системы может потребовать значительных затрат времени и ресурсов для восстановления потерянной в результате инцидента информации и восстановления работоспособности системы. А в некоторых случаях восстановление информации и, соответственно, работоспособности системы становится просто невозможным.

Примером воздействия ошибки в программном обеспечении на работу и безопасность сложной системы может служить катастрофа, произошедшая с первым космическим аппаратом «Mariner-1». Данный аппарат в соответствии с программой «Маринер» должен был стартовать 22 июля 1962 г. и лететь в сторону Венеры [1]. Однако примерно через 5 мин после старта произошла авария, и аппарат был разрушен. В свидетельских показаниях, полученных во время послеполетного расследования инцидента, указывалось, в частности, что в результате ошибки в программном обеспечении управляющие подсистемы аппарата получали неверно интерпретированную информацию. В итоге компьютер аппарата дал серию ошибочных команд на корректировку курса, в итоге аппарат потерял управление и потерпел крушение [2, 3].

Иллюстрацией того, к каким непроизводительным затратам времени приводит разрушение информации может быть инцидент, произошедший в OCLC Research Library Partnership (некоммерческий компьютерный библиотечный сервис и научно-исследовательская организация, цель которых заключается в предоставлении широкого доступа к различной информации в режиме он-лайн) [4]. Эта организация предоставляет доступ к информации с помощью RLIN (компьютерная информаци-

онная сеть библиотек). В этой информационной сети произошла авария, в результате которой был потерян большой объем данных. На восстановление разрушенной информации было потрачено 12 дней кропотливой работы ИТ-специалистов организации [5].

Причины, вызывающие возникновение инцидентов в процессе работы РСОД, приводящих к потере, разрушению или искажению информации и, в итоге, к нарушению безопасности и даже к отказу в работе системы, можно классифицировать следующим образом.

- Сбои и ошибки в работе оборудования вычислительной системы (ВС), на базе которой работает РСОД: сбои в каналах связи, перебои в системах электропитания, сбои в системах хранения информации, сбои в работе вычислительных машин (серверов, рабочих станций).
- Некорректная работа обслуживающего персонала и пользователей системы.
- Потери информации в системах архивации в результате сбоев оборудования.
- Ошибочные данные и потеря информации при некорректной работе программного обеспечения, содержащего ошибки в коде.
- Заражение системы компьютерными вирусами.
- Потери и искажение данных, вызванные несанкционированным доступом в систему.

Для повышения безопасной работы распределенных систем обработки данных вводят различного рода избыточность [5]. В таких системах задача обеспечения сохранности используемых программ и данных решается главным образом применением информационной избыточности (созданием и хранением резервных данных). Кроме того, с этой же целью применяется и аппаратная избыточность (использование резервных устройств хранения данных), а также временная избыточность — использование части ресурсов системы для выполнения процессов дампования (создания копий данных) и восстановления данных в случае их разрушения. Основным методом обеспечения сохранности данных в РСОД при ее функционировании служит метод резервирования, который предполагает использование следующих видов избыточности: идентичной избыточности (создания и хранения копий данных) и/или неидентичной (создания и хранения предыстории данных).

## 1. ПОСТАНОВКА ЗАДАЧИ

В настоящее время в системах обработки данных (СОД) различного класса и назначения применяются два вида резервирования: оперативное и восстановительное.

При оперативном резервировании некоторая часть информационных ресурсов (резервные ко-

пии и/или предыстории массивов данных) РСОД и ее вычислительной мощности используется только для поддержания работоспособности системы в случае, если при решении РСОД текущих задач произошло разрушение используемого ей массива/массивов данных.

В отличие от оперативного резервирования в восстановительном резервировании часть ресурсов системы используется только для восстановления используемых системой рабочих/текущих копий и/или предыстории массивов данных в случае их разрушения.

При оперативном резервировании данных в СОД используются три основные стратегии оперативного резервирования [5, 6]. Первая стратегия применяется для массивов постоянных данных (к этой категории можно отнести относительно редко меняющиеся программные модули или, например, справочники, содержащие статичную информацию разного рода). Вторая и третья стратегии применяются к текущим массивам оперативных данных, содержащим данные, которые подвергаются частым изменениям (добавление и удаление данных, изменение данных, их сортировка), например, таблица с проводками по аналитическим счетам в автоматизированной учетной системе банка или программный модуль на этапе его тестирования и исправления обнаруженных ошибок.

**Стратегия I** заключается в использовании некоторого числа копий основного массива данных. В случае, если основной массив разрушен, то для продолжения решения задачи СОД переходит к использованию первой его копии, если же и она разрушена, то используется следующая копия и т. д.

**Стратегия II** основывается на учете особенностей обновления текущих массивов данных и основывается на том, что вместо копий текущего массива используются его предыстории (предыдущие версии массива и журналы их изменений). В случае разрушения текущего массива запускается специальная программа, которая восстанавливает разрушенный массив на основе его предыдущей версии с использованием записей журнала изменений. Если восстановление закончилось неудачно, то восстановление текущего массива производится из следующей предыстории и т. д.

**Стратегия III** — это смешанная стратегия, которая использует возможности первых двух стратегий, т.е. эта стратегия использует и копии и предыстории текущего массива данных. Причем, сначала в соответствии со стратегией I используются копии массива, а затем, в случае неудачи восстановления из копий, используются предыстории массива в соответствии со стратегией II.

Объединение автономных компьютеров посредством каналов связи позволило получить единую распределенную ВС, обладающую качественно



новыми возможностями по созданию и эксплуатации распределенных СОД по сравнению с возможностями систем, работающих на базе отдельных компьютеров. Применение методов резервирования в таких ВС по сравнению с автономными ЭВМ обладает рядом особенностей [7]:

- при анализе эффективности и выборе оптимальной стратегии резервирования данных в РСОД, построенной на базе ВС, необходимо учитывать параметры самой ВС: топологию сети, задержку сообщений в каналах связи, характеристики производительности отдельных компьютеров, надежность и пропускную способность объединяющих их каналов связи и т. д.;
- в силу территориальной распределенности в вычислительной сети можно использовать один из следующих основных способов хранения резерва РСОД:
  - централизованное, при котором резерв хранится в одном из узлов ВС;
  - децентрализованное, когда резерв распределен по некоторому множеству узлов ВС;
- в случае децентрализованного хранения резерва в РСОД для обработки поступающих в систему запросов можно применять несколько дисциплин обработки. Рассмотрим основные из них (см. рисунок) [7, 8].

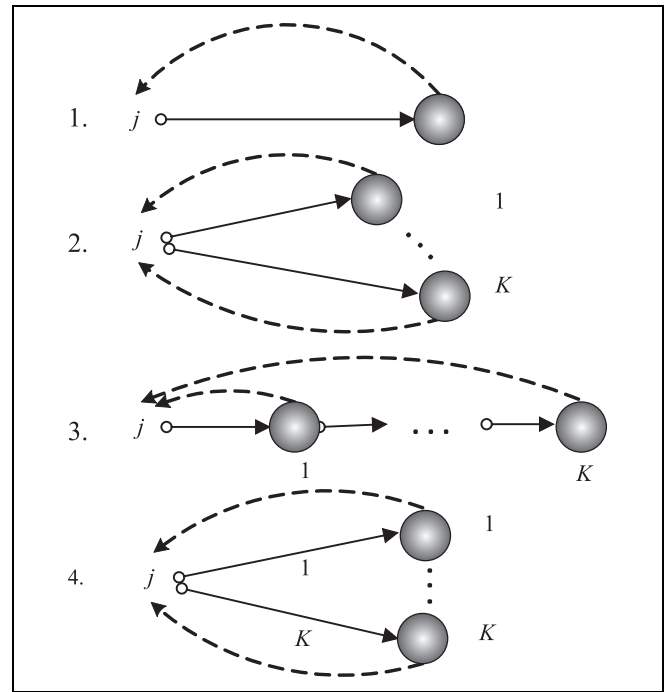
1. Запрос пересылается для обработки в ближайший (согласно некоторому критерию) узел, который хранит резерв массива данных, необходимого для обработки запроса.

2. Запрос пересылается для обработки параллельно в  $K$  узлов с требуемым резервом.

3. Запрос, который изначально был получен в  $j$ -м узле, последовательно пересылается по узлам пути длины  $K$ , начинающегося в  $j$ -м узле и состоящим из узлов, содержащих резерв требуемого массива. Запрос последовательно пересылается по узлам данного пути до тех пор, пока он или не будет успешно обработан в некотором узле либо не будет пройден весь путь из  $K$  узлов.

4. Запрос, полученный в  $j$ -м узле, поочередно отправляется для обработки в  $K$  ближайших узлов, содержащих необходимый резерв. Опрос узлов продолжается до тех пор, пока из очередного опрашиваемого узла не будет успешно получен ответ на запрос либо не будут опрошены все  $K$  узлов без получения требуемого ответа.

Территориальная распределенность ВС обуславливает не только возможность применения различных дисциплин обработки запросов, но и большое разнообразие вариантов размещения резерва по узлам сети, которые обеспечивают различные значения стоимости, времени и надежности обработки поступающих запросов [7]. Эти обстоятельства определяют необходимость решения задачи оптимального размещения оперативного



Основные дисциплины обработки запросов

резерва массивов данных по узлам ВС в соответствии с различными критериями оптимизации. В результате решения этой задачи определяется множество  $I^0 = \{i | y_i = 1\}$  узлов сети, в которых необходимо разместить резерв ( $y_i = 1$ , если в  $i$ -м узле размещен резерв), а также количество копий и/или предыстории массива данных, размещаемых в узлах этого множества  $I^0$ . В качестве критериев оптимизации решения данной задачи можно, например, использовать такие критерии, как максимум вероятности и минимум среднего времени получения ответа на запрос, минимум стоимости обработки запросов и минимум затрат на хранение резерва.

Например, задача поиска размещения оперативного резерва в узлах ВС, обеспечивающего минимум стоимостных затрат  $S(X)$  на обработку поступающих запросов и хранение резерва формулируется таким образом:

$$S(X) \rightarrow \min \tag{1}$$

при ограничениях на время  $t_j$  обработки запросов, вероятность  $\rho_j$  получения ответа на запрос и на размер  $x_j$  резерва в узлах сети:

$$t_j(X) \leq \bar{t}_j, \quad j \in J_3, \tag{2}$$

$$\rho_j(X) \geq \bar{\rho}_j, \quad j \in J_3, \tag{3}$$

$$x_j \leq \bar{x}_j, \quad j \in I^0. \tag{4}$$

Здесь  $J_3$  — это множество индексов узлов ВС, в которых возникают запросы;  $I^0 = \{i | y_i = 1\}$  — множество узлов ВС, в которых размещен резерв;  $X = (x_j)$  — распределение резерва по узлам сети, где  $x_j$  — число копий и/или предысторий, размещенных в узле  $j$ ;  $\bar{x}_j$  — максимальное число копий и/или предысторий массива данных, которые можно разместить в узле  $j$ .

Размер затрат  $S(X)$  для дисциплин 1 и 2 обработки запросов определяется по формуле

$$S(X) = L \sum_{i \in I^0} s_i x_i + \sum_{j=1}^N \sum_{i \in I^0} ZP_{ji}(x_i)(U_j \psi_{ji} + V_j).$$

Здесь  $\psi_{ji} \in \Psi^0$ ; матрица  $\Psi^0 = \|\psi_{ji}^0\|$ , где  $\psi_{ji}^0 \in \{0, 1\}$  и  $\psi_{ji}^0 = 1$ , если запрос из  $j$ -го узла адресуется в  $i$ -й узел,  $s_i$  — стоимость хранения одного бита информации в узле  $i$ ;  $L$  — длина массива данных в битах;  $U_j(V_j)$  — интенсивность информационных запросов (запросов на модификацию), возникающих в узле  $j$ ;  $ZP_{ji}(x_i)$  — размер средних производительных затрат на обработку запроса, посланного из узла  $j$  на обработку в узел  $i$ :

$$ZP_{ji}(x_i) = 2D_{ji} + E_i(x_i)h_i,$$

где  $E_i(x_i)$  — среднее время обработки запроса в узле  $i$ ,  $D_{ji}$  — стоимость передачи запроса/ответа из узла  $j$  в узел  $i$  (предполагается, что  $D_{ji} = D_{ij}$ ),  $h_i$  — стоимость использования ресурсов ЭВМ узла  $i$  в единицу времени.

Значения времени обработки запросов и вероятности получения ответа на запрос определяются по формулам

для дисциплины 1:

$$t_j(X) = 2t_3(1 - \psi_{jj}) + \sum_{k=1}^N E_k(x_k)\Psi_{jk};$$

$$\rho_j(X) = \sum_{k=1}^N r_{jk}P_k(x_k)\Psi_{jk},$$

для дисциплины 2:

$$t_j(X) = \max \left\{ 2t_3 + \max_{\substack{k \in N_j \\ k \neq j}} E_k(x_k); E_j(x_j)\Psi_{jj} \right\},$$

$$\rho_j(X) = 1 - \prod_{k=1}^N [1 - r_{jk}P_k(x_k)r_{kj}\Psi_{jk}],$$

где  $E_k(x_k)$  — среднее время обработки запроса в узле  $k$ ,  $P_k(x_k)$  — вероятность успешной обработки запроса в узле  $k$  при условии размещения в этом узле резерва объемом  $x_k$ .

## 2. РЕШЕНИЕ ЗАДАЧИ

В работе [7] показано, что сформулированная в настоящей статье задача поиска оптимального оперативного резерва принадлежит классу задач целочисленного выпуклого программирования. Для этого в указанной работе доказывается выпуклость ограничений (2)—(4) и целевой функции (1).

Достаточное условие совпадения локального и глобального минимумов целевой функции заключается в выпуклости функций  $S(X)$ ,  $t_j(X)$  и  $\rho_j(X)$ . Данное обстоятельство позволяет применить для решения нашей задачи оптимизации оперативного резервирования методы, применяемые для решения задач распределения ограниченных ресурсов [9].

Аналогично можно показать, что к классу задач целочисленного выпуклого программирования принадлежат и задачи оптимального оперативного резервирования, приведенные ниже и использующие в качестве критериев оптимизации:

— максимум вероятности успешной обработки запросов (задача (5)—(9));

— минимум среднего времени получения ответа на запрос (задача (9)—(12)).

Формулировки данных задач оптимального оперативного резервирования:

Найти:

$$\max R(X) = \max_{j \in J_3} \rho_j(X) \quad (5)$$

при ограничениях:

$$S(X) \leq \bar{S}, \quad (6)$$

$$t_j(X) \leq \bar{t}, \quad j \in J_3, \quad (7)$$

$$x_j \leq \bar{x}_j, \quad j \in I^0. \quad (8)$$

Найти:

$$\min T(X) = \min N_3^{-1} \sum_{j \in J_3} t_j(X), \quad (9)$$

где  $N_3 = |J_3|$ ,

при ограничениях:

$$S(X) \leq \bar{S}, \quad (10)$$

$$\rho_j(X) \geq \bar{\rho}, \quad j \in J_3, \quad (11)$$

$$x_j \leq \bar{x}_j, \quad j \in I^0. \quad (12)$$

Приведенные задачи оптимального оперативного резервирования массивов данных в ВС сформулированы в предположении, что они решаются в два этапа. На первом этапе определяется оптимальное размещение копий массива по узлам ВС  $I^0$  (не более одной копии в каждом узле). Это раз-



мещение получается в результате решения задачи (1)–(4) с помощью известных алгоритмов [10–14].

Если найденное на первом этапе размещение копий не обеспечивает выполнение ограничения (3) задачи, т. е. не обеспечивает требуемую вероятность обработки запросов  $\bar{\rho}$ , то переходят ко второму этапу решения задачи. На втором этапе для каждого узла множества  $I^0$  решается задача определения оптимального объема резерва в отдельном узле ВС. Для ее решения применяются традиционные методы (методы решения задач распределения ограниченных ресурсов) [7].

Решение задачи оптимального оперативного резервирования массивов данных в ВС в два этапа позволяет уменьшить вычислительную сложность задачи, так как вместо одной сложной задачи решаются две более простые задачи. Решение первой из них позволяет определить подмножество узлов ВС, в которых необходимо разместить резерв. А решение второй задачи позволяет определить объем резерва, размещаемого в полученном ранее подмножестве узлов ВС.

Рассмотрим более общую постановку задачи оптимального оперативного резервирования, в которой могут использоваться аналогичные критерии оптимизации.

Рассмотрим такую постановку задачи, например, для дисциплины 1 обработки запросов при использовании в качестве критерия оптимизации минимума стоимости обработки запросов и стоимости хранения резерва. Задача формулируется следующим образом:

найти минимум значения функции

$$S(X) = L \sum_{i \in I^0} s_i x_i + \sum_{j=1}^N V_j \sum_{i/x_i \neq 0} [2D_{ji} + E_i(x_i)h_i] + \sum_{j=1}^N U_j V_{ji} [2D_{ji} + E_i(x_i)h_i]$$

при ограничениях:

$$\rho_j(X) = \rho_{jk} \geq \bar{\rho}, \quad j \in J_3; \quad t_j(X) = t_{jk} \leq \bar{t}, \quad j \in J_3;$$

$$x_j \in \{0, 1, \dots, \bar{x}_j\}, \quad j = \overline{1, N},$$

где  $k$  — индекс узла сети с оперативным резервом, для которого справедливы условия:

$$2D_{jk} + E_k(x_k)h_k = \min_{i \in N_j} [2D_{ji} + E_i(x_i)];$$

$$N_j = \{i/\rho_{ji} \leq \bar{\rho}, t_{ji} \leq \bar{t}\};$$

$$\rho_{ji} = r_{ji}P_i(x_i)r_{ij} \text{ при } i \neq j \text{ и } \rho_{ji} = P_i(x_i) \text{ при } i = j;$$

$$t_{ji} = 2t_3 + E_i(x_i) \text{ при } i \neq j \text{ и } t_{ji} = E_i(x_i) \text{ при } i = j,$$

где  $E_k(x_k)$  — среднее время обработки одного запроса в узле  $k$ ;  $t_3$  — среднее время задержки пере-

дачи сообщений в каналах связи сети;  $h_k$  — стоимость использования ресурсов компьютера узла  $k$  в единицу времени;  $D_{jk}$  — стоимость передачи сообщения (запроса/ответа на запрос) из узла  $j$  в узел  $k$  (полагаем, что  $D_{jk} = D_{kj}$ ).

Аналогично формулируются и другие задачи оптимального оперативного резервирования, использующие в качестве критерия оптимизации минимум среднего времени получения ответа на запрос или максимум вероятности успешной обработки запросов.

### 3. РЕЗЕРВИРОВАНИЕ ВЗАИМОСВЯЗАННЫХ МАССИВОВ ДАННЫХ

Задачи оптимального оперативного резервирования (см. § 1 и 2) были сформулированы в предположении, что резервируемые в РСОД массивы данных используются независимо друг от друга. Это позволило ставить и решать задачи оптимального резервирования отдельно и независимо для каждого массива данных. Однако чаще всего на практике в СОД разного класса и назначения, в том числе и в РСОД, построенных на базе ВС, решаются задачи и обрабатываются запросы, работа с которыми требует использования некоторого подмножества массивов данных. Причем, как правило, в данное подмножество входят массивы данных, взаимосвязанных между собой (ссылками или алгоритмически). В частности, обработка одного запроса к распределенной базе данных (РБД) также может потребовать доступ к данным из нескольких связанных друг с другом массивов данных, которые могут быть размещены в разных узлах ВС, территориально удаленных друг от друга.

По этой причине возникает необходимость оптимизации резервирования массивов данных, используемых в РСОД, с учетом их взаимосвязи друг с другом и других параметров ВС (таких, как топология сети, трафик запросов различного типа, надежность функционирования каналов связи и отдельных ЭВМ сети и др.). Покажем, что данные задачи можно сформулировать, используя приведенные выше результаты.

Рассмотрим работу РСОД, построенную на основе ВС, состоящей из  $N$  узлов. В системе используется  $M$  массивов данных и решается  $J$  различных задач. Каждая из задач системы может использовать несколько различных массивов и может решаться в нескольких различных узлах сети. Обозначим через  $W_j$  то число решений  $j$ -й задачи,  $j = \overline{1, J}$ , которое производится за некий интервал времени (например, час, сутки, неделя, месяц, квартал и т. п.).

Введем переменную  $\Psi_{jn} \in (0, 1)$  такую, что  $\Psi_{jn} = 1$ , если  $j$ -я задача решается в  $n$ -м узле сети; обозначим  $U_{jm}$  — среднее число информационных запросов  $j$ -й задачи к  $m$ -му массиву данных, а  $V_{jm}$  — среднее число запросов на модификацию  $m$ -го массива данных, возникающих при решении  $j$ -й задачи;  $x_{nm}$  — объем резерва  $m$ -го массива, размещенного в  $n$ -м узле сети,  $L_m$  — размер  $m$ -го массива данных в байтах.

Условимся, что задача в системе решена успешно только в случае, если получены ответы на все запросы, возникающие в процессе ее решения (запросы на получение и модификацию данных) ко

всем требуемым массивам данных. Тогда вероятность  $P_{nj}(X)$  решения  $j$ -й задачи в  $n$ -м узле при распределении резерва  $X = \|x_{nm}\|$

$$P_{nj}(X) = \prod_{m=1}^M [\rho_{nm}(X)]^{U_{jm}} \beta_{nm}^{V_{jm}},$$

где  $\rho_{nm}(X)$  — вероятность получения ответа на информационный запрос к  $m$ -му массиву, выданный в  $n$ -м узле сети;  $\beta_{nm}(X)$  — это вероятность успешной обработки выданного в  $n$ -м узле сети запроса на обновление  $m$ -го массива данных. Значение данной вероятности зависит от того, какая используется дисциплина обработки запроса, и опреде-

**Таблица вероятностных временных и стоимостных характеристик дисциплин обработки запросов**

Дисциплина	Характеристика	
1	$\rho_j = \sum_{k=1}^N r_{jk} P_k(x_k) r_{kj} \Psi_{jk}$	$Z = \sum_{j=1}^N \sum_{k=1}^N (U_j \Psi_{jk} + V_j y_k) Z_{jk}(x_k)$
	$t_j = 2t_3(1 - \Psi_{jj}) + \sum_{k=1}^N E_k(x_k) \Psi_{jk}$	$ZP = \sum_{j=1}^N \sum_{k=1}^N (U_j \Psi_{jk} + V_j y_k) ZP_{jk}(x_k)$
2	$\rho_j = 1 - \prod_{k=1}^N [1 - r_{jk} P_k(x_k) r_{kj} \Psi_{jk}]$	$Z = \sum_{j=1}^N \sum_{k=1}^N (U_j \Psi_{jk} + V_j y_k) Z_{jk}(x_k)$
	$t_j = \max \left\{ 2t_3 + \max_{k \in N_j^*, k \neq j} E_k(x_k); E_j(x_{kj}) \Psi_{jj} \right\}$	$ZP = \sum_{j=1}^N \sum_{k=1}^N (U_j \Psi_{jk} + V_j y_k) ZP_{jk}(x_k)$
3	$\rho_j = \sum_{n=1}^K r_{jn} P_n(x_n) r_{nj} \prod_{i=2}^n (1 - P_{j_{i-1}}(x_{j_{i-1}}) r_{j_{i-1}j})$	$Z = \sum_{j=1}^N \sum_{k=1}^N (U_j + V_j y_k) Z_{jk}(x_k)$
	$t_j = (K + 1)t_3 + \sum_{k=1}^N E_k(x_k) \Psi_{jk} - B; B = t_3 \text{ при } n(j, j) = 1 \text{ или } n(j, j) = K \text{ и } 0 \text{ в остальных случаях}$	$ZP = \sum_{j=1}^N \sum_{k=1}^N (U_j + V_j y_k) ZP_{jk}(x_k)$
4	$\rho_j = 1 - \prod_{k=1}^N [1 - r_{jk} P_k(x_k) r_{kj} \Psi_{jk}]$	$Z = \sum_{j=1}^N \sum_{k=1}^N \left\{ U_j \prod_{i=1}^{n(j,k)-1} [1 - r_{ji} P_i(x_i) r_{ji}] + V_j y_k \right\} Z_{jk}(x_k),$ $j_i \in N_j; i = \overline{1, K}$
	$t_j = 2t_3(K - \Psi_{jj}) + \sum_{k=1}^N E_k(x_k) \Psi_{jk}$	$ZP = \sum_{j=1}^N \sum_{k=1}^N \left\{ U_j \prod_{i=1}^{n(j,k)-1} [1 - r_{ji} P_i(x_i) r_{ji}] + V_j y_k \right\} ZP_{jk}(x_k),$ $j_i \in N_j; i = \overline{1, K}$
Обозначения: $Z_{jk}(x_k) = 2(1 - r_{jk})D_{jk} + [1 - P_k(x_k)]E_k(x_k)h_k$ ; $ZP_{jk}(x_k) = 2D_{jk} + E_k(x_k)h_k$ ; $U_{ik} = \sum_{j=1}^N \Psi_{ji} \Psi_{jk} U_{ik}^{(j)}$ при $n(j, k) = n(j, i) + 1$ и 0 в остальных случаях; $U_{ik}^{(j)} = U_i \prod_{\alpha=1}^{n(j,i)-1} r_{j_{\alpha-1}j_{\alpha}} [1 - P_{j_{\alpha}}(x_{j_{\alpha}})]$ , $j_{\alpha} = j, j_{\alpha} \in N_j$		



ляется по соответствующей формуле, приведенной в таблице [7]. Предположим, что значение этой вероятности равно вероятности события, заключающегося в обновлении по крайней мере одного из массивов данных, размещенного в одном из узлов сети с резервом, и последующего получения сообщения об успешном обновлении этого массива. С учетом сделанного предположения вероятность успешной обработки запроса на обновление массива данных

$$\beta_{nm}(X) = 1 - \prod_{i=1}^N [1 - r_{ni} P_i(x_{im}) r_{in}^i],$$

где  $P_i(x_{im})$  — вероятность успешной обработки запроса в узле  $i$ , в котором размещен резерв объемом  $x_{im}$ ;  $r_{ni}$  — надежность канала связи между  $n$ -м и  $i$ -м узлом ВС (будем считать, что  $r_{ni} = r_{in}$ ).

Обозначим через  $T_{nj}(X)$  суммарное время ожидания ответов на все запросы, выданные при решении  $j$ -й задачи в  $n$ -м узле ВС. Оно определяется по формуле

$$T_{nj}(X) = \sum_{m=1}^M \{U_{jm} t'_{nm}(X) + U_{jm} t''_{nm}(X)\},$$

где  $t'_{nm}(X)$  — время получения ответа на запрос к  $m$ -му массиву, выданный в  $n$ -м узле (его значение вычисляется по соответствующей формуле из таблицы);  $t''_{nm}(X)$  — время обработки одного запроса на модификацию данных в  $m$ -м массиве, который был выдан в узле  $n$ . Значение данного времени также определяется по формуле из таблицы для дисциплины 2 обработки запросов при  $N_j = \{i/x_{im} \neq 0\}$ .

В таблице приведены формулы для расчета характеристик дисциплин обработки запросов:  $\rho_j$  — вероятности получения ответа на запрос, выданный в  $j$ -м узле ВС;  $t_j$  — времени обработки запроса, выданного в узле  $j$ ;  $ZP$  — размера средних производительных затрат на обработку запросов;  $Z$  — размера средних непроизводительных затрат на обработку запросов.

Стоимость  $S(X)$  функционирования РСОД в течение рассматриваемого нами периода времени будет складываться из стоимости решения всех задач системы и стоимости хранения резерва в узлах сети и, таким образом,

$$S(X) = \sum_{m=1}^M \sum_{j=1}^J W_j \sum_{n=1}^N \varphi_{jn} \left( V_{jm} \sum_{i/x_{im} \neq 0} ZP_{ni}(x_{im}) + U_{jm} \min_{i/x_{im} \neq 0} ZP_{ni}(x_{im}) \right) + L_m \sum_{n=1}^N s_n x_{nm},$$

где  $ZP_{ni}(x_{im}) = 2D_{ni} + E_i(x_{im})h_i$ ,  $s_n$  — стоимость хранения одного бита информации в  $n$ -м узле в течение рассматриваемого периода времени.

Сформулируем задачу оптимального оперативного резервирования  $M$  взаимосвязанных массивов данных при использовании в качестве критерия минимума стоимостных затрат на функционирование системы  $S(X)$ .

Необходимо найти такое оптимальное размещение резерва используемых в РСОД массивов данных по узлам ВС, которое обеспечит минимальные стоимостные затраты на функционирование системы. Найденное решение задачи должно соответствовать ограничениям на объем резерва, размещаемого в каждом из узлов ВС, и должно обеспечивать затраты времени на ожидание ответов на запросы не более заданного лимита.

Таким образом, найти

$$S(X) \rightarrow \min$$

при ограничениях:

$$P_{nj}(X) \geq \bar{P}_j, \quad T_{nj}(X) \leq \bar{T}_j, \quad \sum_{m=1}^M x_{nm} L_m \leq Q_n, \\ j = \overline{1, J}; \quad n = \overline{1, N}.$$

Аналогично формулируются задачи поиска оптимального размещения в ВС резерва взаимосвязанных массивов, используемых в РСОД, при других критериях оптимальности решения. Например, при использовании такого критерия оптимальности размещения резерва, как максимум вероятности успешного решения всех задач системы за заданный период времени, задача формулируется следующим образом:

$$P(X) = \prod_{j=1}^J \prod_{n=1}^N [P_{nj}(X)]^{W_j} \rightarrow \max$$

при ограничениях

$$S(X) \leq \bar{S}, \quad T_{nj}(X) \leq \bar{T}_j, \quad \sum_{m=1}^M x_{nm} L_m \leq Q_n, \\ j = \overline{1, J}, \quad n = \overline{1, N}.$$

#### 4. РЕЗЕРВИРОВАНИЕ ИНФОРМАЦИИ В РАСПРЕДЕЛЕННЫХ БАЗАХ ДАННЫХ

Рассмотрим ситуацию, когда необходимо найти оптимальное размещение резерва взаимосвязанных массивов для распределенной базы данных (РБД), в адрес которой пользователями ВС формируются поток информационных запросов и поток запросов на модификацию данных.

Обозначим:  $J$  — множество номеров информационных запросов к массивам, обрабатываемым средствами РБД;  $J^o$  — множество номеров различных запросов на модификацию данных в массивах РБД;  $M_j$  — множество номеров массивов, используемых при обработке  $j$ -го информационного за-

проса,  $j \in J$ ;  $M_j^o$  — множество номеров массивов, которые модифицируются при выполнении одного запроса  $j^o \in J^o$ . В каждый фиксированный интервал времени в адрес РБД в  $n$ -м узле сети в среднем формируется  $U_{nj}$  запросов  $j$ -го типа (информационный запрос к массиву данных) и  $V_{nj^o}$  запросов  $j^o$ -го типа (модификация массива данных).

Вероятность  $R_n(X)$  получения ответов на все запросы, выданные в  $n$ -м узле за рассматриваемый интервал времени (т. е. надежность обслуживания пользователей  $n$ -го узла сети, генерирующих поток запросов к РБД) определяется по формуле:

$$R_n(X) = \prod_{j \in J} \left[ \prod_{m \in M_j} \rho_{nm}(X) \right]^{U_{nj}} \prod_{j^o \in J^o} \left[ \prod_{m \in M_j^o} \beta_{nm}(X) \right]^{V_{nj^o}}.$$

Всеми пользователями  $n$ -го узла ВС, работающими с РБД, за рассматриваемый интервал времени будет затрачено  $T_n(X)$  единиц времени на ожидание ответов на выданные ими запросы:

$$T_n(X) = \sum_{j \in J} U_{nj} \sum_{m \in M_j} t'_{nm}(X) + \sum_{j^o \in J^o} V_{nj^o} \sum_{m \in M_j^o} t''_{nm}(X).$$

Стоимость функционирования РБД за рассматриваемый интервал времени (затраты на хранение резерва и обработку запросов)

$$S(X) = \sum_{n=1}^N \left[ \sum_{j \in J} U_{nj} \sum_{m \in M_j} \min_{i/x_{im} \neq 0} ZP_{ni}(x_{im}) + \sum_{j^o \in J^o} V_{nj^o} \sum_{m \in M_j^o} \sum_{i/x_{im} \neq 0} ZP_{ni}(x_{im}) \right] + \sum_{m=1}^M L_m \sum_{n=1}^N S_n x_{nm}.$$

Полученные формулы позволяют сформулировать задачу оптимизации резервирования массивов данных в РБД по критерию максимума надежности обслуживания пользователей базы данных — необходимо обеспечить максимальную вероятность обработки всех запросов пользователей к РБД, сформированных за рассматриваемый интервал времени, при ограничениях на объем резерва, размещаемого в каждом из узлов сети, стоимость функционирования РБД, общее время ожидания ответов на запросы, выданные пользователями одного узла за рассматриваемый интервал времени:

$$R(X) = \prod_{n=1}^N R_n(X) \rightarrow \max$$

при ограничениях  $T_n(X) \leq \bar{T}_n$ ,  $S(X) \leq \bar{S}$ ,

$$\sum_{m=1}^M x_{nm} L_m \leq Q_n, \quad n = \overline{1, N}.$$

Аналогично формулируются и задачи оптимального резервирования массивов данных в РБД с другими критериями оптимизации: минимум стоимости функционирования РБД за заданный интервал времени; минимум суммарного времени ожидания ответов на запросы пользователей РБД, сформированные ими за заданный интервал времени.

## 5. РАЗМЕЩЕНИЕ РЕЗЕРВА ИНФОРМАЦИИ С УЧЕТОМ ИНТЕНСИВНОСТИ ТРАФИКА ЗАПРОСОВ

Оптимальное распределение оперативного резерва массивов данных по узлам ВС находится в непосредственной зависимости от трафика запросов к резервируемым массивам, возникающих в узлах сети при решении задач и формируемых пользователями сети. Выполним качественный анализ влияния трафика запросов, возникающих в узлах ВС, на оптимальное распределение оперативного резерва массивов данных. Анализ выполним на примере однородной полносвязной ВС в смысле размера стоимостных затрат на хранение резерва и обработку запросов —  $S(Y)$ . Будем считать, что запросы на модификацию массива данных адресуются одновременно во все узлы сети с резервом, а информационные запросы к массиву обрабатываются в соответствии с дисциплиной 1. Затраты будем вычислять по формуле

$$S(Y) = [h + 2D + h^*(T_Y - T + 2t_3)] \sum_{j=1}^N (U_j + KV_j) + sK - (2D + h^*2t_3) \sum_{j \in Y} (U_j + V_j), \quad (13)$$

где  $D$  — стоимость передачи запроса/ответа по каналам связи ВС;  $h$  — стоимость обработки одного запроса в узле с резервом;  $s$  — стоимость хранения резерва в одном из узлов;  $Y$  — множество индексов узлов сети, по которым распределен резерв (распределение резерва по ВС);  $K = |Y|$ ;  $T_Y$  — среднее время обработки запроса в одном из узлов сети при распределении резерва  $Y$ ;  $T$  — заданная верхняя граница времени получения ответа на запрос;  $h^*$  — стоимость (штраф) превышения временем получения ответа величины  $T$ .

Сделаем предположения:

— информационные запросы, возникающие в узлах без резерва, равномерно распределяются между всеми  $K$  узлами сети с резервом;

— запросы, поступающие в каждый из узлов с резервом, образуют пуассоновский поток с пара-





метром  $\lambda_y$ ; время обработки запросов распределено по показательному закону;

— для каждого узла сети  $j = \overline{1, N}$  выполняются равенства:  $U_j = U$ ;  $V_j = V$ .

Основываясь на сделанных предположениях, получим, что  $T_Y = t/(1 - \lambda_y t)$ , где  $t$  — среднее время обработки запроса, а  $\lambda_y = NV + NK^{-1}U$ .

Предположим, что размещение резерва  $Y'$  мы получили на основе распределения  $Y$  путем размещения в узле  $i$  дополнительного резерва. Определим, при каких соотношениях между значениями  $U$  и  $V$  справедливо неравенство

$$S(Y') > S(Y), \quad (14)$$

в котором  $Y' = Y \cup \{i\}$ ; ( $K' = |Y'| = K + 1$ ).

При сделанных предположениях из формулы (13) следует, что

$$S(Y') = S(Y) - N(U + KV)(T_Y - T_{Y'})h^* + S - (U + V)(2D + h^*2t_3) + NV[h + 2D + h^*(T_Y - T) + h^*2t_3], \quad (15)$$

а неравенство (14) будет эквивалентно неравенству

$$V[Nh + 2D(N - 1) + Nh^*(T_Y - T) + h^*2t_3(N - 1) - NK(T_Y - T_{Y'})h^*] > U[Nh^*(T_Y - T_{Y'}) + 2D + h^*2t_3] - S. \quad (16)$$

Достаточным условием выполнения неравенства (16) служит соотношение между значениями  $U$  и  $V$ :

$$V > U[Nh^*(T_Y - T_{Y'}) + 2D + h^*2t_3] \times [Nh + 2D(N - 1) + Nh^*(T_Y - T) + h^*2t_3(N - 1) - NK(T_Y - T_{Y'})h^*]^{-1}. \quad (17)$$

Рассмотрим частные случаи.

1. Значение  $S(Y)$  зависит главным образом от затрат на использование каналов связи, т. е.  $D \gg h^*$ ,  $D \gg h$ . В этом случае соотношение (17) приводится к виду

$$V \geq U(N - 1)^{-1}. \quad (18)$$

2. Превышение временем получения ответа на запрос величины  $T$  несущественно, т. е.  $h^* \ll h$ ,  $h^* \ll D$ .

Соотношение (17) выполняется при условии

$$V \geq U(N - 1)(h/2D + 1)^{-1}. \quad (19)$$

3. Учитывается главным образом время получения ответа на запрос, т. е.  $h^* \gg D$ ,  $h^* \gg h$ . В этом случае выражение (17) эквивалентно соотношению

$$V \geq U[N(T_Y - T_{Y'}) + 2t_3][N(T_Y - T) + (N - 1)2t_3 - N_K(T_Y - T_{Y'})]^{-1}. \quad (20)$$

Полученные качественные результаты (при сделанных предположениях) сформулируем в виде утверждения,

**Утверждение 1.** В однородной полностью связной вычислительной сети при выполнении соотношения (17) (или соотношений (18)—(20)) соответственно для рассмотренных случаев 1—3 наименьшие стоимостные затраты на хранение резерва и на обработку запросов достигаются при размещении резерва в одном узле сети.

Из выражения (15) следует, что неравенство  $S(Y') < S(Y)$  справедливо при выполнении соотношения

$$U > V\{Nh + 2D(N - 1) + h^*[N(T_Y - T) - 2t_3(N - 1) - NK(T_Y - T_{Y'})]\}B^{-1} + SB^{-1}, \quad (21)$$

где  $B = Nh^*(T_Y - T_{Y'}) + 2D + h^*2t_3$ .

Рассмотрим частные случаи:

1) если  $D \gg h^*$ ,  $D \gg h$ ,  $D \gg S$ , то соотношение (21) эквивалентно неравенству

$$U > V(N - 1); \quad (22)$$

2) при  $h^* \ll h$ ,  $h^* \ll D$  и  $h^* \ll S$  соотношение (21) приводится к виду

$$U > V(N - 1) + (S + NhV)(2D)^{-1}; \quad (23)$$

3) для  $h^* \gg D$  и  $h^* \gg h$  получим:

$$U > V[N(T_Y - T) + 2t_3(N - 1) - NK(T_Y - T_{Y'})][N(T_Y - T_{Y'}) + 2t_3]^{-1}. \quad (24)$$

Обобщим полученные качественные результаты в виде утверждения, справедливого при сделанных предположениях.

**Утверждение 2.** В однородной полностью связной вычислительной сети при выполнении соотношения (21) (или соотношений (22)—(24)) соответственно для рассмотренных случаев 1—3 наименьшие стоимостные затраты на хранение резерва и обработку запросов обеспечивает размещение резерва во всех узлах, в которых возникают запросы.

**Пример.** Покажем на примере решения задачи оптимального оперативного резервирования массивов данных в ВС то, как влияет на оптимальное размещение резерва изменение значений параметров самой сети. Рассмотрим результат решения задачи оптимального резервирования на примере сегмента однородной полностью связной ВС из пяти узлов. Параметры узлов рассматриваемого сегмента сети (в силу однородности сети индексы узлов опущены):

$$\theta = 1,5; \quad \tau = 0,5; \quad U = 20; \quad V = 7; \quad q = 0,01; \quad S = 2,0;$$

$$r^2 = 0,985; \quad 2D = 9,0; \quad h = 3,0; \quad N = 5.$$

В качестве критерия задачи примем минимум стоимостных затрат на обработку запросов и хранение резерва.

Для рассматриваемого нами примера сегмента сети с указанными параметрами выполняется соотношение (18). Следовательно, справедливо и утверждение 1, из которого следует, что при заданных параметрах сети резерв целесообразно разместить только в одном из узлов.

Проверка данного утверждения путем решения задачи оптимального оперативного резервирования методом полного перебора показывает, что для рассматриваемого фрагмента сети оптимальными являются пять равнозначных (в силу однородности сети) распределений резерва: (20000), (02000), (00200) и (00002). Эти распределения резерва обеспечивают одинаковое значение целевой функции  $S(Y) = 1792,07$ .

Если же в параметрах рассматриваемого сегмента сети произойдут изменения:  $V = 3$ ,  $h = 1,5$ , то в этом случае будет выполняться соотношение (23) и будет уже справедливо утверждение 2. В соответствии с этим утверждением при измененных параметрах ВС наиболее целесообразно размещение резерва в каждом из узлов сегмента, в котором возникают запросы. Решение задачи оптимального оперативного резервирования методом полного перебора при новых параметрах фрагмента сети доказывает справедливость утверждения 2 и дает нам одно оптимальное распределение оперативного резерва — (11111), которое обеспечивает следующее значение целевой функции  $S(Y) = 1337,49$ .

### ЗАКЛЮЧЕНИЕ

Рассмотрены три основные стратегии оперативного резервирования массивов постоянных и текущих оперативных данных и указаны особенности применения этих стратегий в системах обработки данных, работающих в распределенных вычислительных сетях. Описаны различные варианты хранения резерва в вычислительной сети и несколько возможных дисциплин обработки информационных запросов и запросов на модификацию данных массивов. Сформулирована в общем виде задача распределения информационного резерва по узлам вычислительной сети, оптимального по критерию минимума суммарных затрат на обработку поступающих запросов и хранение резерва в узлах сети. Также сформулированы задачи оптимального резервирования массивов данных в вычислительных сетях по критериям максимума вероятности успешной обработки запроса к распределенным системам обработки данных и минимума среднего времени получения ответа на запрос. Даны формулировки задач оптимального оперативного резервирования массивов данных в вычислительных сетях с учетом различного рода взаимосвязей между резервируемыми массивами и задач оптимального резервирования в распределенных базах данных. Приведены методы решения поставленных задач оптимизации. На примере однородной полносвязной сети ЭВМ выполнен качественный анализ влияния интенсивностей запросов, возникающих в отдельных узлах сети, на

распределение оперативного резерва массивов данных. Изложенные в статье результаты имеют практическую ценность и могут быть использованы при проектировании распределенных систем обработки данных в целях повышения безопасности их функционирования.

### ЛИТЕРАТУРА

1. NASA publication SP-480, Far Travelers -The Exploring Machines, Oran W. Nicks, 1985. — URL: <http://history.nasa.gov/SP-480/ch3.htm> (дата обращения 14.07.2016).
2. NASA Space Science Data Coordinated Archive. Mariner 1. NSSDCA ID: MARINI. — URL: <http://nssdc.gsfc.nasa.gov/nmc/spacecraftDisplay.do?id=MARINI> (дата обращения 23.08.2016).
3. Mariner Venus Final Project Report (NASA SP-59, 1965). — P. 87. — URL: <http://history.nasa.gov/SP-4407/ETUv5.pdf> (дата обращения 14.07.2016).
4. History of the OCLC Research Library Partnership. — URL: <http://www.oclc.org/research/partnership/history.html> (дата обращения 14.07.2016).
5. Информационная безопасность систем организационного управления. Теоретические основы: в 2 т. / Н.А. Кузнецов, В.В. Кульба, Е.А. Микрин и др.; отв.ред. Н.А. Кузнецов, В.В. Кульба; Ин-т проблем передачи информации РАН. — М.: Наука, 2006.
6. Кульба В.В., Мамиконов А.Г., Шелков А.Б. Резервирование программных модулей и информационных массивов в АСУ // Автоматика и телемеханика. — 1980. — № 8. — С. 133—141.
7. Кульба В.В., Сомов С.К., Шелков А.Б. Резервирование данных в сетях ЭВМ. — Казань: Изд-во Казанского гос. ун-та, 1987. — 175 с.
8. Сомов С.К. Резервирование программных модулей и информационных массивов в сетях ЭВМ: дисс. канд. техн. наук. — М.: ИПУ РАН, 1983. — 217 с.
9. Гуринов Л.С., Дымарский Я.С., Меркулов А.Д. Задачи и методы оптимального распределения ресурсов. — М.: Сов. радио, 1968. — 463 с.
10. Machmoud S., Riordon J.S. Optimal Allocation of Resources in Distributed Information networks // ACM Trans. on Database Systems. — 1976. — Vol. 1, N 4. — P. 66—78.
11. Chu W.W. File Allocation in an Multiple Computer System // IEEE Trans. on Computers. — 1969. — Vol. C-18, N 10. — P. 885—889.
12. Casey R.G. Allocations of copies of a file in an Information Network // AFIPS Conference Proceedings. — 1972. — Vol. 40. — P. 617—625.
13. Akoka J. Design of optimal distributed database systems. Distributed Data Bases // Proc. Int. Symp. — 1980. — P. 229—246.
14. Ramamoorthy C.V., Wah B.W. Data management in distributed data bases // AFIPS Conference Proceedings. — 1978. — Vol. 48. — P. 667—680.
15. Клейнрок Л. Вычислительные системы с очередями. — М.: Мир, 1979. — 600 с.

Статья представлена к публикации членом редколлегии В.В. Кульбой.

**Микрин Евгений Анатольевич** — академик РАН, ген. конструктор, ПАО «Ракетно-космическая корпорация «Энергия» им. С.П. Королева», г. Королев, [✉ Eugeny.Mikrin@rsce.ru](mailto:Eugeny.Mikrin@rsce.ru),

**Сомов Сергей Константинович** — канд. техн. наук, вед. инженер, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, [✉ ssomov2016@ipu.ru](mailto:ssomov2016@ipu.ru).