

АНАЛИЗ МЕТОДОВ ОПЕРАТИВНОГО РЕЗЕРВИРОВАНИЯ ИНФОРМАЦИИ В ВЫЧИСЛИТЕЛЬНЫХ СЕТЯХ

Е.А. Микрин, С.К. Сомов

Рассмотрены особенности и предложены формализованные модели методов оперативного резервирования информации в системах обработки данных, работающих на базе вычислительных сетей. Выполнен сравнительный анализ централизованного и децентрализованного вариантов хранения резерва. Для стационарного и переходного режимов функционирования ВС проведен сравнительный анализ трех основных стратегий оперативного резервирования информации.

Ключевые слова: система обработки данных, вычислительная сеть, стратегия оперативного резервирования, централизованное и децентрализованное хранение резерва.

ВВЕДЕНИЕ

Конец прошлого и начало нынешнего века ознаменовался бурным развитием локальных и глобальных вычислительных сетей (ВС) и созданных на их основе распределенных систем обработки данных (РСОД) различного масштаба и назначения.

Быстрое развитие технологий, качественный и количественный рост вычислительных сетей и созданных на их основе информационных систем ставят в разряд первоочередных проблем решение задач и разработку методов повышения надежности и безопасности функционирования РСОД, работающих в рамках ВС. Повышение надежности РСОД обеспечивается повышением надежности и безопасности применения программного и информационного обеспечения локальных компьютеров и их объединением в единую сеть посредством надежных высокоскоростных каналов связи.

Эффективным способом повышения надежности РСОД является также и оперативное резервирование массивов данных в вычислительных сетях [1, 2]. К массивам данных можно отнести, например, программные модули, информационные массивы, справочники, таблицы баз данных.

Решение задачи обеспечения сохранности информации известно и широко применяется в сис-

темах обработки данных разного масштаба и назначения. Суть решения заключается в использовании дополнительных ресурсов вычислительной техники для резервирования массивов данных, что позволяет значительно сократить воздействие разрушающих факторов на информацию и тем самым повысить безопасность функционирования системы обработки данных (СОД).

Проблема обеспечения сохранности информации в вычислительных сетях начала изучаться довольно давно. Например, статьи по данной проблеме западными учеными были опубликованы в конце 1970-х — начале 1980-х годов [3—5]. Примерно в это же время работы по проблеме повышения достоверности и сохранности информации, в том числе и в вычислительных сетях, начали активно исследоваться и в Институте проблем управления РАН [6—11]. С развитием возможностей вычислительной техники, средств связи и методов разработки программного обеспечения актуальность данной тематики не уменьшилась, работы по данной тематике ведутся как в России [12, 13], так и за рубежом [14—17].

Эти работы в ИПУ РАН не прекращались, ведутся они и в настоящее время [18, 19]. Были развиты новые направления в области повышения безопасности работы СОД и получены значимые результаты. Например, были начаты исследования по таким новым направлениям в области обеспе-

чения безопасности СОД, как восстановительное резервирование, оперативное резервирование, резервирование и восстановление данных в распределенных вычислительных системах с учетом разных дисциплин обработки запросов к данным, резервирование информации в РСОД реального времени, обеспечение информационной безопасности систем организационного управления, повышение безопасности распределенных баз данных [18–21]. В рамках этих направлений в ИПУ РАН было предложено использовать модель случайного блуждания частицы по целочисленным точкам действительной прямой, что было сделано впервые, в том числе и для решения задач резервирования информации в вычислительных сетях.

1. ПОСТАНОВКА ЗАДАЧИ И ОСОБЕННОСТИ РЕЗЕРВИРОВАНИЯ ДАННЫХ В ВЫЧИСЛИТЕЛЬНЫХ СЕТЯХ

В настоящей статье рассматриваются формализованные модели и методы оперативного резервирования информации в РСОД, работающих на базе корпоративных, специальных или глобальных вычислительных сетей с учетом особенностей резервирования массивов данных в ВС и различных дисциплин обработки запросов к информации. Цель статьи состоит в анализе эффективности методов оперативного резервирования информации в вычислительных сетях.

При работе СОД существует вероятность разрушения используемых массивов данных вследствие различных разрушающих факторов [7, 8]. Это, как правило, приводит к появлению различных ошибок в результатах работы системы, к невозможности решения одной или некоторого множества задач и даже к полной потере работоспособности системы. Восстановление разрушенных массивов данных и работоспособности системы может потребовать значительных затрат ресурсов и времени.

Обеспечение сохранности информации — это многоплановая и комплексная задача, и оперативное резервирование массивов данных служит одним из системных методов ее решения. Он заключается в том, что некоторая часть ресурсов РСОД (резервные копии и/или предыстории массивов данных) и некоторая доля вычислительных мощностей используются только для поддержания работоспособности системы в случае, если при функционировании РСОД произошло разрушение используемой информации.

Компьютеры, объединенные каналами связи и системным программным обеспечением в одну вычислительную сеть, обладают качественно иными свойствами и возможностями по сравнению с отдельными, автономными компьютерами. Поэтому

применение методов резервирования информации для повышения безопасности РСОД, работающих на основе ВС, обладают рядом особенностей по сравнению с СОД, работающими на базе автономных компьютеров. [1, 11].

В частности, при анализе эффективности и выборе оптимальной стратегии резервирования информации в вычислительных сетях необходимо учитывать топологию сети, надежность, стоимостные и временные характеристики каналов связи и компьютеров. Необходимо также рассматривать разные варианты распределения резерва по узлам сети: децентрализованное (резерв распределен по некоторому подмножеству узлов ВС), централизованное (резерв хранится в одном из узлов сети) и динамическое.

Кроме того, децентрализованное хранение резерва предоставляет возможность применения нескольких дисциплин обработки запросов, возникающих в ВС. Различают четыре основные дисциплины обработки запросов [1, 10].

1. Запрос, сформированный в некотором узле ВС, пересылается для обработки в ближайший (согласно выбранному критерию) узел ВС, в котором размещен резерв требуемого для обработки запроса массив данных (рис. 1, *a*).

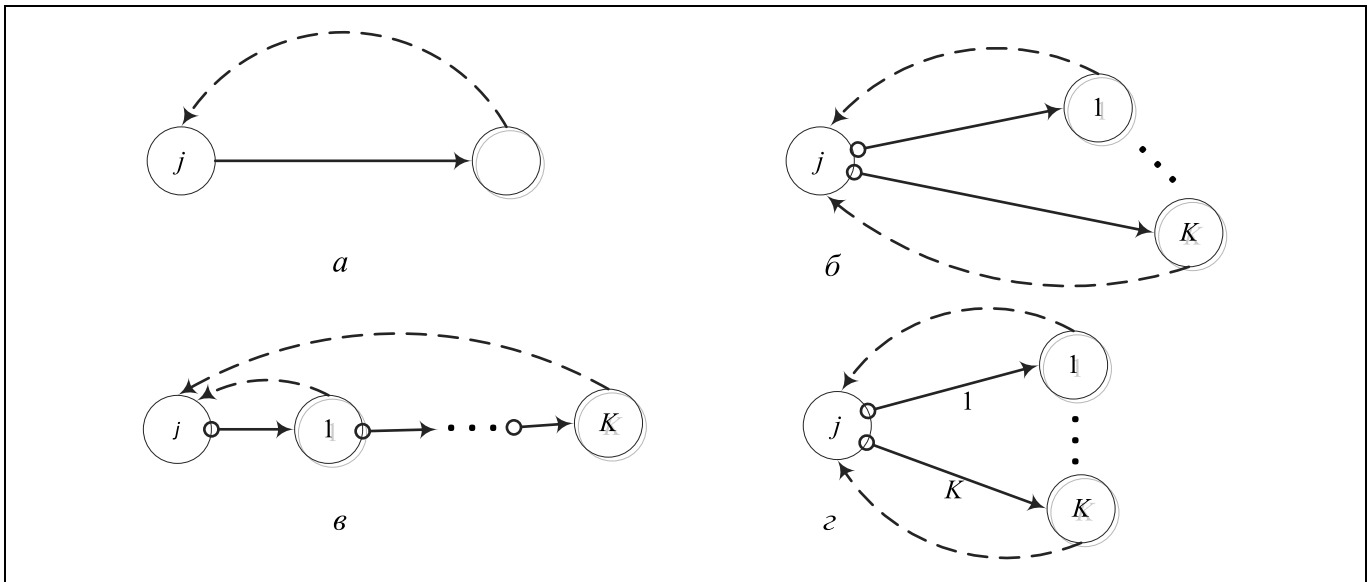
2. Запрос одновременно пересылается для обработки в K ближайших узлов ВС с необходимым резервом (рис. 1, *b*).

3. Запрос, сформированный в j -м узле сети, пересылается последовательно по узлам пути, начинающимся в этом узле и имеющим длину K узлов с резервом необходимого массива данных. Запрос пересылается по данному пути до тех пор, пока в некотором узле k , $k = \overline{1, K}$, этого пути запрос будет успешно обработан либо не будет пройден весь путь, а запрос так и не будет обработан (рис. 1, *в*).

4. Из j -го узла, в котором сформирован запрос, запрос последовательно пересылается для обработки в K ближайших узлов сети с необходимым резервом. Процесс опроса узлов продолжается до момента, когда либо из очередного узла будет получен ответ, либо будут опрошены все K узлов, а ответ на запрос так и не был получен (рис. 1, *г*).

Анализ стратегий резервирования и выбор оптимальной стратегии для случая централизованного хранения резерва осуществляется традиционными методами [7, 8]. При этом необходимо учитывать временные характеристики ВС стоимость и надежность использования каналов связи и компьютеров.

Территориальная распределенность ВС, наличие в них большого числа узлов и каналов связи с разными характеристиками надежности, производительности и стоимости их использования пре-



Основные дисциплины обработки запросов:

a, б, в и г — дисциплины 1, 2, 3 и 4 соответственно

доставляют большое число различных вариантов размещения резерва в сети, из которых необходимо выбрать один оптимальный вариант. Дополнительную сложность данной задаче придает возможность применения различных дисциплин обработки запросов. Эти обстоятельства обуславливают необходимость формулировки задач поиска оптимального размещения резерва в узлах ВС и разработки методов их решения. В качестве критериев оптимизации при постановке таких задач можно применять различные критерии: минимум стоимости функционирования РСОД (минимум затрат на обработку запросов и хранение резерва массивов данных системы), максимум вероятности обработки запроса, минимум среднего времени получения ответа на запрос.

2. СТРАТЕГИИ РЕЗЕРВИРОВАНИЯ МАССИВОВ ДАННЫХ

В СОД различных классов и назначения в настоящее время применяются три основные стратегии резервирования массивов данных, которые были разработаны в ИПУ РАН с учетом особенностей использования и обновления массивов данных [1, 7, 8, 21].

Стратегия I. Согласно данной стратегии резерв информации представляет собой некоторое число копий массива данных. Если при обработке запроса используемый при этом основной массив разрушается, то для продолжения обработки запроса используется первая копия массива, если и она

разрушается, то обработка запроса продолжается при использовании следующей копии, и т. д. Процесс использования копий заканчивается либо при успешной обработке запроса, либо при разрушении всех копий резерва (запрос не обработан).

Стратегия II. В качестве резерва массива данных вместо копий массива используется некоторое число его предысторий (предыстория массива данных это предыдущая его полная копия и журнал последующих изменений данных в массиве). При разрушении основного массива во время обработки запроса он сначала восстанавливается специальной программой обновления из его последней предыстории и затем продолжается обработка запроса. Если восстановление было неудачным, то делается попытка восстановления массива из следующей, более ранней его предыстории, и т. д.

Стратегия III. Это смешанная стратегия, согласно которой в качестве резерва используются и копии массива данных, и его предыстории. Согласно данной стратегии в случае разрушения основного массива для продолжения обработки запроса сначала в соответствии со стратегией I используются копии массива, а если запрос не обработан и резерв копий массива разрушен, то делается попытка восстановления основного массива с использованием резерва предысторий согласно стратегии II.

В работе [21] приведены формулы для расчета вероятностных и временных характеристик данных стратегий: вероятности $P(x)$ и среднего времени $E(x)$ обработки запроса к массиву данных, имеющего резерв объемом x копий и/или предысторий.

3. АНАЛИЗ ЭФФЕКТИВНОСТИ СТРАТЕГИЙ ОПЕРАТИВНОГО РЕЗЕРВИРОВАНИЯ В ВЫЧИСЛИТЕЛЬНЫХ СЕТЯХ

Сравним эффективность трех стратегий резервирования, когда они применяются в двух вариантах хранения резерва — централизованном и децентрализованном. Сравним будем по показателям эффективности резервирования: ρ_j — надежность обработки запроса, т. е. вероятность получения ответа на запрос, выданный в j -м узле сети; $q_j = (1 - \rho_j)$ — вероятность того, что запрос из узла j не будет обработан по какой-либо причине, и ответ на этот запрос не будет получен в узле j ; ZP — размер средних затрат на использование ресурсов ВС для обработки запросов при условии их успешной обработки; Z — размер средних непроизводительных затрат на использование ресурсов ВС для обработки запросов; t_j — время получения ответа на запрос, выданный в j -м узле.

Эффективности стратегий будем сравнивать при условиях:

- распределение резерва по узлам ВС одинаково для всех стратегий;
- во всех узлах сети применяется одна и та же стратегия резервирования;
- среднее время задержки сообщений в сети постоянно и равно t_3 ;
- запросы в системе делятся на информационные запросы к массивам данных и запросы на модификацию данных;
- запросы на модификацию данных одновременно пересылаются во все узлы сети, в которых размещен резерв модифицируемых данных;
- для обработки поступающих информационных запросов применяется одна из четырех дисциплин обработки, описанных в § 1.

Рассмотрим применение дисциплины 4 обработки запросов при децентрализованном варианте хранения резерва и получим аналитические выражения для расчета показателей эффективности стратегий резервирования.

Согласно дисциплине 4 обработки запросов, запрос, возникший в j -м узле ВС, будет поочередно пересылаться для обработки в K ближайших узлов, содержащих резерв данных, необходимых для обработки запроса ($K = |N_j|$). Здесь N_j — множество индексов узлов с резервом, ближайших к узлу j . При этом запрос пересылается в очередной узел множества N_j , если он не был успешно обработан в предыдущем узле множества.

Введем переменную $\psi_{jk} \in \{0, 1\}$ такую, что $\psi_{jk} = 1$, если $k \in N_j$. Тогда при дисциплине 4 обработки запросов значение вероятности

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

Здесь: $[1 - r_{jk} P_k(x_k) r_{kj} \psi_{jk}]$ — вероятность события, заключающегося в том, что из очередного обрабатываемого узла k ответ на запрос, посланный из узла j , не был получен; r_{jk} — надежность канала связи между j -м и k -м узлом; $P_k(x_k)$ — вероятность успешной обработки запроса в узле k при размещении в этом узле резерва, имеющего объем x_k .

При заданных условиях значение t_j времени получения ответа на запрос, выданный в узле j , будет определяться по формуле

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

где $E_k(x_k)$ — среднее время обработки запроса в узле k .

Значение $Z_{jk}(x_k)$ средних непроизводительных затрат ресурсов ВС, затраченных на неуспешную обработку одного запроса, посланного из j -го узла для обработки в k -й узел, будет рассчитываться по формуле:

$$Z_{jk}(x_k) = 2(1 - r_{jk})D_{jk} + [1 - P_k(x_k)]E_k(x_k)h_k.$$

Здесь: h_k — стоимость использования ресурсов компьютера узла k за единицу времени; D_{jk} — стоимость передачи запроса или ответа из j -го узла в узел k (исходим из предположения, что $D_{jk} = D_{kj}$).

Значение Z средних непроизводительных затрат ресурсов ВС при обработке всех запросов, возникающих в сети за единицу времени, с учетом обработки запросов в соответствии с рассматриваемой дисциплиной 4 определяется как

$$Z = \sum_{j=1}^N \sum_{k=1}^N \left\{ U_j \prod_{i=1}^{n(j,k)-1} [1 - r_{ji} P_i(x_i) r_{ij}] + V_j y_k \right\} Z_{jk}(x_k),$$

где $n(j, k)$ — порядковый номер узла k в множестве N_j , $j_i \in N_j$, $i = \overline{1, K}$, $y_k \in \{0, 1\}$; $y_k = 1$, если в узле k размещен требуемый резерв; величины U_j и V_j — число возникающих в узле j за единицу времени (интенсивность) информационных запросов и число запросов на модификацию данных соответственно.

Расчет значения ZP средних производительных затрат ресурсов ВС, используемых для обработки



запросов, проводится по предыдущей формуле, в которой необходимо величину $Z_{jk}(x_k)$ заменить на величину

$$ZP_{jk}(x_k) = 2D_{jk} + E_k(x_k)h_k.$$

Формулы для расчета показателей эффективности стратегий резервирования при других дисциплинах обработки запросов и вариантов хранения резерва в узлах ВС получаются аналогичным образом.

В табл. 1 и 2 приведены аналитические выражения для расчета показателей эффективности рассматриваемых в статье стратегий резервирования,

соответствующих децентрализованному и централизованному вариантам хранения резерва.

Аналитические выражения для расчета характеристик стратегий резервирования применительно к двум частным случаям динамического хранения резерва приведены в табл. 3.

Динамический резерв в обоих частных случаях перемещается по N_1 узлам множества J_1 , $|J_1| = N_1$, узлов сети и представляет собой M блоков по X копий и/или предысторий массива данных. Все блоки динамического резерва перемещаются одновременно с частотой β перемещений в единицу времени. В первом частном случае маршрут пере-

Таблица 1

Вероятностные, временные и стоимостные характеристики стратегий резервирования при децентрализованном хранении резерва

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

Таблица 2

Вероятностные, временные и стоимостные характеристики стратегий резервирования при централизованном хранении резерва

Характеристика	
$\rho_j = \sum_{k=1}^N r_{jk} P_k(x_k) r_{kj} y_{jk}$	$Z = \sum_{j=1}^N \sum_{k=1}^N (U_j + V_j) Z_{jk}(x_k)$
$t_j = 2t_3(1 - y_j) + \sum_{k=1}^N E_k(x_k) y_k$	$ZP = \sum_{j=1}^N \sum_{k=1}^N (U_j + V_j) ZP_{jk}(x_k)$

мещения блоков фиксированный, во втором случае блоки динамического резерва перемещаются случайным образом. Для второго случая блок резерва из j -го узла перемещается в i -й с вероятностью, которая определяется как $\psi_{ji} = 1/N_1, j, i = \overline{1, N_1}$.

В формулах табл. 3 приняты обозначения: J_1 — множество индексов узлов сети, содержащих резерв; переменная y_i , которая равна 1 если $i \in J_1$ и 0 в противном случае; d_{ji} — стоимость пересылки бита информации из j -го в i -й узел; переменная

$\bar{x}_{ji} = 1$, если блок динамического резерва из j -го узла перемещается в узел i , иначе переменная равна 0; L — длина в битах копии или предыстории массива данных.

Сравним эффективность рассматриваемых вариантов резервирования по надежности обработки запросов, т. е. по значению вероятности ρ_j получения ответа на запрос, выданный в j -м узле сети.

В работах [1, 10] показано, что при использовании стратегий I—III резервирования в ВС и выполнении условий $x_i > 0, 0 < q_i < 1/2, i \in N_j$, для вероятности успешной обработки запроса в одном узле справедливо соотношение:

$$P_i^I(x_i) > P_i^{II}(x_i) > P_i^{III}(x_i).$$

Отсюда следует, что для всех вариантов применения стратегий I—III резервирования для вероятности ρ_j получения ответа на запрос, выданный в j -м узле сети, справедливо соотношение [1]:

$$\rho_j^I > \rho_j^{II} > \rho_j^{III}.$$

Данный результат сформулируем как

Утверждение 1. Применение в вычислительных сетях стратегии I резервирования массивов данных

Таблица 3

Вероятностные, временные и стоимостные характеристики стратегий резервирования (частные случаи динамического хранения резерва)

№ случая	Характеристика	
1	$\rho_j = 1 - \prod_{k=1}^N [1 - r_{jk} M N_1^{-1} P_k(X) r_{kj} y_k]$ $t_j = \max\{2t_3 + \max_{k/y_k=1; k \neq j} E_k(X); E_j(X) y_j\}$	$Z = \sum_{j=1}^N \sum_{k=1}^N (U_j + V_j) M N_1^{-1} Z_{jk}(X) y_k + M \beta X L N_1^{-1} \sum_{j=1}^N \sum_{k=1}^N \bar{X}_{jk} (1 - r_{jk}) d_{jk}$ $ZP = \sum_{j=1}^N \sum_{k=1}^N (U_j + V_j) M N_1^{-1} ZP_{jk}(X) y_k + M \beta X L N_1^{-1} \sum_{j=1}^N \sum_{k=1}^N \bar{X}_{jk} d_{jk}$
2	$\rho_j = 1 - \prod_{k=1}^N [1 - r_{jk} \bar{P}_k r_{kj} y_k]$ $t_j = \max\{2t_3 + \max_{k/y_k=1; k \neq j} \bar{E}_k(X); E_j y_j\}$	$Z = \sum_{j=1}^N \sum_{k=1}^N (U_j + V_j) y_k \{ [1 - (1 - N_1^{-1})]^{M2} (1 - r_{jk}) D_{jk} + \bar{B}_k \} +$ $+ M \beta X L N_1^{-2} \sum_{j \in J_1} \sum_{k \in J_1} (1 - r_{jk}) d_{jk}$ $ZP = \sum_{j=1}^N \sum_{k=1}^N (U_j + V_j) y_k \{ [1 - (1 - N_1^{-1})]^{M2} D_{jk} + \bar{E}_k h_k \} +$ $+ M \beta X L N_1^{-2} \sum_{j \in J_1} \sum_{k \in J_1} d_{jk}$
Обозначения: $\bar{P}_k = N_1^{-M} \sum_{m=1}^M P_k(mX) C_M^m (N_1 - 1)^{M-m}$; $\bar{B}_k = N_1^{-M} \sum_{m=1}^M [1 - P_k(mX)] E_k(mX) C_M^m (N_1 - 1)^{M-m}$; $\bar{E}_k = N_1^{-M} \sum_{m=1}^M E_k(mX) C_M^m (N_1 - 1)^{M-m}$		



(хранение копий массивов) по сравнению со стратегиями II и III обеспечивает наибольшую вероятность получения ответа на запрос. ♦

Далее сравним стратегии резервирования по значению времени t_j получения ответа на запрос.

Сначала рассмотрим стационарный режим функционирования ВС, при котором $q_k \rightarrow 0$, $k \in J_1$, где q_k — вероятность разрушения копии/предыстории массива данных в узле k за единичный интервал времени θ_k ее использования, J_1 — множество индексов узлов, с размещенным в них резервом ($|J_1| = N_1$).

Для обработки запроса в одном узле i в стационарном режиме требуется в среднем $E_i(x_i)$ времени, для которого справедливо следующее соотношение [10, 11]:

$$E_i^{\text{II}}(x_i) < E_i^{\text{III}}(x_i) < E_i^{\text{I}}(x_i). \quad (1)$$

Отсюда следует, что для рассматриваемых нами вариантов применения стратегий резервирования массивов данных справедливо неравенство

$$t_i^{\text{II}} < t_i^{\text{III}} < t_i^{\text{I}}.$$

Полученный результат сформулируем как

Утверждение 2. Стратегия II обеспечивает наилучшее время получения ответа на запрос по сравнению со стратегиями I и III при работе сети в стационарном режиме, т. е. при $q_i \rightarrow 0$ и любых θ_i и τ_i (здесь τ_i — время создания одной копии). ♦

Теперь перейдем к рассмотрению переходного режима функционирования ВС, при котором вероятность разрушения массива данных значительна, т. е. $q_i \rightarrow 1/2$.

В переходном режиме функционирования сети для среднего времени $E_i(x_i)$ обработки запроса в одном узле i при $\tau_i > \theta_i$ выполняется соотношение (2) [6, 10]:

$$E_i^{\text{III}}(x_i, l_i > 2) < E_i^{\text{II}}(x_i) < E_i^{\text{III}}(x_i, l_i = 2) < E_i^{\text{I}}(x_i), \quad (2)$$

где l_i — число предыстории массива данных, используемых в стратегии резервирования III.

Соотношение (2) позволяет доказать

Утверждение 3. В переходном режиме функционирования сети при $q_i \rightarrow 1/2$, $\tau_i > \theta_i$, $i \in N_j$, для всех рассматриваемых вариантов резервирования для времени t_i справедливо соотношение:

$$t_i^{\text{III}}(l_i > 2) < t_i^{\text{II}} < t_i^{\text{III}}(l_i = 2) < t_i^{\text{I}}. \quad \blacklozenge$$

Из утверждения 3 следует, что при рассмотренных выше условиях стратегия III является наилучшей по значению времени получения ответа.

Если параметры ВС таковы, что $q_i \rightarrow 1/2$, $\theta_i \gg \tau_i$, $i \in N_j$, то для величины $E_i(x_i)$ и времени t_i будут справедливы следующие соотношения:

$$E_i^{\text{I}}(x_i) < E_i^{\text{III}}(x_i) < E_i^{\text{II}}(x_i); \quad t_i^{\text{I}} < t_i^{\text{III}} < t_i^{\text{II}}. \quad (3)$$

Следовательно, справедливо

Утверждение 4. В переходном режиме функционирования сети при значении параметров $q_i \rightarrow 1/2$, $\theta_i \gg \tau_i$, стратегия I обеспечивает наименьшее время ответа на запрос.

Размер \bar{Z}_i средних непроизводительных затрат ресурсов компьютера i -го узла сети при обработке одного запроса вычисляется по формуле [10]:

$$\bar{Z}_i(x) = [1 - P_i(x)]E_i(x)h_i.$$

Если проанализировать данную формулу [10], то мы увидим, что при значениях параметров сети $q_i \rightarrow 1/2$, $\theta_i \gg \tau_i$, выполняется соотношение:

$$\bar{Z}_i^{\text{I}}(x) < \bar{Z}_i^{\text{III}}(x) < \bar{Z}_i^{\text{II}}(x).$$

Данное соотношение позволяет сделать вывод о том, для непроизводительных затрат Z при рассматриваемых нами трех стратегиях резервирования массивов данных справедливо соотношение:

$$Z^{\text{I}} < Z^{\text{III}} < Z^{\text{II}}.$$

Таким образом, справедливо

Утверждение 5. Применение в вычислительных сетях стратегии I оперативного резервирования массивов данных при параметрах сети $q_i \rightarrow 1/2$, $\theta_i \gg \tau_i$, обеспечивает наименьшие непроизводительные затраты на использование ее ресурсов при обработке запросов.

В стационарном режиме работы ВС, когда $q_i \rightarrow 0$, значения среднего времени $E_i(x_i)$ обработки запроса в узле i находятся в соотношении (1). Тогда очевидно

Утверждение 6. В стационарном режиме функционирования вычислительной сети ($q_i \rightarrow 0$, $i = \overline{1, N}$), для стратегий I, II и III справедливо соотношение:

$$ZP^{\text{II}} < ZP^{\text{III}} < ZP^{\text{I}}. \quad \blacklozenge$$

Таким образом, в стационарном режиме функционирования ВС наилучшей стратегией по размеру производительных затрат ZP на обработку поступающих запросов является стратегия II (хранение предыстории массивов данных).

В переходном режиме функционирования ВС (при $q_i \rightarrow 1/2$, $\theta_i \gg \tau_i$) значение среднего времени обработки запроса $E_j(x_j)$ в узле сети удовлетворяет соотношению (3). Из этого следует [10], что для всех рассматриваемых дисциплин обработки запросов как для децентрализованного, так и для централизованного варианта резервирования справедливо соотношение:

$$ZP^I < ZP^{III} < ZP^{II}, \quad (4)$$

т. е. доказано.

Утверждение 7. В переходном режиме функционирования вычислительной сети (при $q_i \rightarrow 1/2$, $\theta_i \gg \tau_i$) наименьшие затраты ресурсов сети на обработку запросов обеспечивает стратегия I оперативного резервирования.

Для переходного режима функционирования ВС, при значениях параметров ВС $q_i \rightarrow 1/2$, $\tau_i > \theta_i$, для среднего времени $E_j(x_j)$ обработки запроса в одном узле справедливо соотношение (2). Анализ этого соотношения показывает [10], что для всех рассматриваемых дисциплин обработки запросов и вариантов применения стратегий резервирования выполняется соотношение:

$$ZP_i^{III} (l_i > 2) < ZP_i^{II} < ZP_i^{III} (l_i = 2) < ZP_i^I,$$

т. е. наименьшие затраты ресурсов сети на обработку запросов для централизованного, динамического и децентрализованного вариантов резервирования в ВС с параметрами $q_i \rightarrow 1/2$, $\tau_i > \theta_i$ обеспечивает стратегия III оперативного резервирования.

4. ПРИМЕР

Рассмотрим применение описанных моделей для расчета и анализа показателей ρ_j , t_j и ZP и стратегий резервирования на примере фрагмента вычислительной сети, состоящего из девяти узлов, и резервирования одного массива данных. Параметры рассматриваемого фрагмента сети представлены в табл. 4.

В итоге решения задачи оптимального резервирования по критерию минимума затрат на обработку запросов получено оптимальное распределение резерва по узлам сети, которое представлено в табл. 5.

Время задержки сообщений в каналах связи сети и надежность соответствующих каналов имеют значения: $t_3 = 1,0$; $r_{jk} = r_{kj}$ для $i, k = \overline{1,9}$; $r_{jk} = 1,0$ при $j = k$; $r_{12} = r_{23} = r_{13} = r_{45} = r_{65} = r_{78} = 0,975$ и $r_{jk} = 0,95$ для остальных i, k .

Запросы обрабатываются в соответствии с дисциплиной 4 при $K = 2$ (максимальное число опрашиваемых узлов) и со следующими множествами N_j узлов с резер-

вом, в которых могут обрабатываться запросы к массиву, возникающие в узлах $j = \overline{1,9}$:

$$N_1 = N_3 = \{1, 2\}; \quad N_2 = \{2, 1\}; \quad N_4 = N_6 = \{1, 5\};$$

$$N_5 = \{5, 1\}; \quad N_7 = N_8 = \{7, 1\}; \quad N_9 = \{9, 1\}.$$

Значения показателей эффективности стратегий резервирования, рассчитанные при заданных параметрах сети в соответствии с аналитическими выражениями, приведенными в § 3, представлены в табл. 6 и 7.

Таблица 4

Параметры фрагмента сети

Параметр сети	Номер узла сети i								
	1	2	3	4	5	6	7	8	9
U_j	60	40	10	10	20	15	30	10	40
V_j	6	4	2	2	4	3	4	2	5
q_j	0,05	0,1	—	—	0,1	—	0,1	—	0,1
τ_j	4	10	—	—	10	—	8	—	8
θ_j	2	5	—	—	5	—	4	—	4
h_j	2,0	1,0	—	—	1,0	—	1,2	—	1,2
D_{1j}	0	0,04	0,1	0,5	0,4	0,5	3,4	3,4	2,0
D_{2j}	0,04	0	0,08	0,5	0,4	0,5	3,4	3,4	2,0
D_{5j}	0,4	0,4	0,4	0,06	0	0,06	3,3	3,3	1,9
D_{7j}	3,4	3,4	3,4	3,3	3,3	3,3	0	0,06	4,3
D_{9j}	2,0	2,0	2,0	1,9	1,9	1,9	4,3	4,3	0

Таблица 5

Оптимальное распределение резерва

Номер узла сети i	1	2	3	4	5	6	7	8	9
Объем размещенного резерва x_i	3	3	0	0	3	0	3	0	3

Таблица 6

Вероятности обработки запросов

Стратегия	ρ_1	ρ_2	ρ_3	ρ_4, ρ_6, ρ_8	ρ_5, ρ_7, ρ_9
I	0,9999938	0,9999506	0,9975095	0,9950877	0,9999023
II	0,9999931	0,9999397	0,9974983	0,9950666	0,9998808
III	0,9999934	0,9999456	0,9975042	0,9950781	0,9998925

Таблица 7

Времена обработки запросов и производительные затраты

Стратегия	t_1, t_2, t_5	t_3, t_4, t_6	t_7, t_9	t_8	ZP
I	37,6	39,7	32,2	34,5	9916,18
II	10,4	12,4	9,2	11,2	2772,0
III	23,4	25,7	20,6	22,6	6175,3



ЗАКЛЮЧЕНИЕ

Рассмотрен эффективный метод повышения надежности функционирования распределенных систем обработки данных, заключающийся в применении информационной избыточности. Для уменьшения влияния различных разрушающих факторов, снижающих сохранность информации в распределенной системе, предложено использовать резерв имеющихся в системе массивов данных в виде их копий и/или предысторий. Описаны три стратегии резервирования массивов данных в узлах сети, получены аналитические выражения для расчета их вероятностных, временных и стоимостных характеристик. Рассмотрены варианты децентрализованного, централизованного и динамического хранения резерва, для которых выполнен сравнительный анализ эффективности стратегий по вероятности успешной обработки запроса к данным, среднему времени получения ответа на запрос и средним стоимостным затратам на обработку запросов. Учитывалась возможность применения четырех дисциплин обработки запросов в распределенной системе. Приведен пример расчета значений характеристик стратегий для фрагмента вычислительной сети с помощью предложенной модели резервирования массивов данных.

ЛИТЕРАТУРА

- Кульба В.В., Сомов С.К., Шелков А.Б. Резервирование данных в сетях ЭВМ. — Казань: Изд-во Казанского ун-та, 1987. — 175 с.
- Микрин Е.А., Сомов С.К. Задача синтеза оптимальной СОД РВ с резервированием информации и структурной избыточностью // Тр. XVI междунар. конф. «Проблемы управления безопасностью сложных систем». — М., 2008. — С. 137–141.
- Maohmoud 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.
- Garcia-Molina H. Reliability Issues for Fully Replicated Distributed Databases // Computer. — 1982. — Vol. 15, N 9. — P. 34–42.
- Coffman E.G., Gelenbe E., Plateau B. Optimization of the number of copies in a distributed data base // IEEE Trans. of Software Eng. — 1981. — Vol. 7, N 1. — P. 78–84.
- Кульба В.В. Анализ стратегий резервирования информационных массивов в АСУ // Методы и модели планирования и управления в дискретных производственных системах: Сб. тр. / ИПУ РАН. — 1977. — Вып. 14. — С. 20–32.
- Кульба В.В., Шелков А.Б., Пелихов В.П. Стратегии резервирования информационных массивов // Построение автоматизированных систем обработки данных: Сб. тр. / ИПУ РАН. — 1978. — Вып. 16. — С. 26–42.
- 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.
- Кульба В.В., Мамиконов А.Г., Пелихов В.П. Методы повышения достоверности и сохранности информации в АСУ: Обзор // Автоматика и телемеханика. — 1985. — № 2. — С. 5–33.
- Сомов С.К. Резервирование программных модулей и информационных массивов в сетях ЭВМ: дис. ... канд. техн. наук. — М.: ИПУ РАН, 1983. — 217 с.
- Мамиконов А.Г., Кульба В.В., Сомов С.К. Анализ стратегий резервирования программных модулей и информационных массивов в сетях ЭВМ // Автоматика и телемеханика. — 1984. — № 2. — С. 149–159.
- Микрин Е.А., Сомов С.К. Оптимальное оперативное резервирование информации в системах обработки данных на базе вычислительных сетей // Проблемы управления. — 2016. — № 5. — С. 47–56.
- Микрин Е.А., Сомов С.К. Оптимизация резервирования информации в распределенных системах обработки данных реального времени // Проблемы управления. — 2016. — № 6. — С. 47–52.
- Abdalla H.I. A synchronized design technique for efficient data distribution // Computers in Human Behavior. — 2014. — Vol. 30. — P. 427–435.
- Mansouri N. Adaptive data replication strategy in cloud computing for performance improvement // Frontiers of Computer Science (print). — 2016. — Vol. 10, N 5. — P. 925–935.
- Sahoo J., Salahuddin M.A., Gliho R. A Survey on Replica Server Placement Algorithms for Content Delivery Networks. — IEEE Communications Surveys & Tutorials. — 2016. — 30 p. — URL: <https://arxiv.org/ftp/arxiv/papers/1611/1611.01729.pdf> (дата обращения 2.05.2017).
- Singh A., Kahlon K.S. Non-replicated dynamic data allocation in distributed database system // Intern. Journal of Computer Science and Network Security. — 2009. — Vol. 9, N 9.
- Информационная безопасность систем организационного управления (теоретические основы): в 2 т. Т. 2 / отв. ред. Н.А. Кузнецов, В.В. Кульба. — М.: ИППИ РАН, 2006. — 437 с.
- Микрин Е.А., Шелков А.Б., Павельев В.В. Методы восстановления данных в распределенных автоматизированных системах: науч. изд. — М.: ИПУ РАН, 2009. — 68 с.
- Информационное обеспечение систем организационного управления (теоретические основы): в 3 ч. Ч. 2. Методы анализа и проектирования информационных систем / Е.А. Микрин и др.; под ред. Е.А. Микрина, В.В. Кульбы. — М.: Физматлит, 2011. — 496 с.
- Информационное обеспечение систем организационного управления (теоретические основы): в 3-х ч. Ч. 3. Методы повышения качества информационного обеспечения систем организационного управления / Е.А. Микрин и др.; под ред. Е.А. Микрина, В.В. Кульбы. — М.: Физматлит, 2012. — 528 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, ✉ Eugeney.Mikrin@rsce.ru,

Сомов Сергей Константинович — канд. техн. наук, ст. науч. сотрудник, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, ✉ ssomov2016@ipu.ru.