В.В. Кульба, С.К. Сомов

Репликация, резервирование и схемы восстановления информации в ненадежных распределенных системах

Статья посвящена проблеме повышения надежности распределенных систем обработки данных, функционирующих на базе ненадежных компьютерных сетей. Рассматривается ситуация, когда в дополнение к размещению в узлах компьютерной сети реплик и резерва массивов данных используются методы восстановления информации в случае ее разрушения. Анализируются условия, при которых целесообразно использование в качестве восстановительного резерва или неразрушенного оперативного резерва ближайшего узла или архива магнитных носителей. Рассмотрена задача определения оптимальной схемы восстановления разрушенных реплик и резерва массивов данных.

Ключевые слова: распределенные системы обработки данных, репликация и резервирование массивов данных, восстановительное резервирование.

На основе глобальных компьютерных сетей разрабатывается и эксплуатируется большее количество распределенных систем обработки данных (далее – РСОД) различного назначения. В рамках таких систем каналами связи объединяются компьютеры, рассредоточенные на расстоянии сотен, а иногда и тысяч километров. В отличие от систем обработки данных, работающих на основе локальных компьютерных сетей, глобальные РСОД рассчитаны на поддержку работы огромного количества пользователей, часто используют уже существующие, не очень качественные и не слишком скоростные каналы связи. Такие РСОД являются крупномасштабными, сложными системами, основанными на взаимодействии огромного числа

программных, аппаратных и информационных компонент. При этом практически каждая из компонент имеет ненулевую вероятность отказа. На надежность работы РСОД влияет множество факторов, которые могут привести к потере, разрушению или искажению информации и в итоге к нарушению безопасности, простою или к отказу самой системы[1].

Важными задачами, которые необходимо решить при проектировании РСОД, являются следующие две: обеспечение высокой надежности работы системы и требуемого уровня ее производительности. Известным и широко используемым методом решения этих задач является метод репликации по узлам компьютерной сети массивов данных, используемых в системе[2].

Метод заключается в том, что несколько реплик (копий) одного массива данных распределяются по узлам компьютерной сети таким образом, чтобы максимально приблизить данные к их потребителям (пользователям или приложениям). Тем самым сокращается время, затрачиваемое на обработку запросов к реплицируемому массиву данных. Другое достоинство метода заключается в том, что он обеспечивает и более высокую надежность работы РСОД, так как при отказе одного узла с репликой массива данных запросы переадресуются в другие работоспособные узлы с репликами данного массива. Идентичность содержимого всех реплик одного массива поддерживается механизмом репликации (тиражирование изменений данных по всем репликам) в синхронном режиме или в режиме асинхронной репликации.

Однако в ненадежных компьютерных сетях существует вероятность того, что будет разрушена не одна, а несколько реплик массива данных, что может привести к значительному, неприемлемому уменьшению производительности системы и уменьшению надежности ее работы.

Для существенного уменьшения вероятности потери информации вследствие разрушения реплик предлагается использовать два метода: метод резервирования массивов данных и метод восстановительного резервирования.

Согласно первому методу, в узлах ненадежной сети предлагается размещать не только сами реплики массива данных, но и дополнительный оперативный резерв (далее – ОР) из копий и/или предысторий этого массива. В этом случае, если при обработке запроса реплика разрушается, вместо повторной попытки обработки запроса в другом узле сети с репликой для продолжения обработки запроса используется размещенный в этом же узле резерв копий и/или предысторий массива данных.

Для создания резерва данных в узлах РСОД с репликами используются три стратегии организации оперативного резервирования[3].

Стратегия 1. Создается резерв из некоторого количества копий реплики. При разрушении реплики используется первая копия, при ее разрушении – вторая и так далее. Применяется для резервирования массивов постоянных данных.

Стратегия 2. Используется для резервирования переменных массивов данных. В узле сети создается резерв из предысторий массива данных – предыдущих поколений массива и массивов изменений (журналов транзакций). При разрушении реплики происходит ее восстановление специальной программой из первой предыстории. При разрушении предыстории она восстанавливается из следующей предыстории и так далее.

Стратегия 3. Смешанная стратегия использует как копии, так и предыстории реплик. Причем вначале используются копии в соответствии со стратегий 1, а при их разрушении – предыстории, как в стратегии 2.

Размещение в узлах сети с репликами оперативного резерва существенно повышает вероятность успешной обработки запросов, однако остается ненулевая вероятность разрушения не только самой реплики, но ее резерва. В этом случае для восстановления работоспособности узла используется метод восстановительного резервирования.

Данный метод заключается в том, что для восстановления разрушенного ОР и реплик используется специальный восстановительный резерв (далее – ВР) из копий и/или предысторий, который используется для целей восстановления разрушенного ОР и реплик. На практике используют два типа восстановительного резервирования[4]:

– в первом случае в качестве ВР используется неразрушенный ОР узла сети, ближайшего к узлу с разрушенной репликой и оперативным резервом. Близость узла определяется согласно некоторому критерию близости узлов РСОД;
– второй тип предполагает, что для восстановления разрушенных данных в качестве ВР используется специальный резерв данных – архив магнитных носителей (далее АМН), который предназначен для длительного и надежного хранения массивов данных и используется исключительно для обработки запросов на восстановление разрушенных данных. АМН может располагаться в одном узле или децентрализованно в нескольких узлах компьютерной сети.

Для восстановления разрушенных данных используются две стратегии восстановления[5]: стратегия В-1 и стратегия В-2. Согласс-

но стратегии В-1 при помощи ВР последовательно получаются все копии разрушенного ОР. Согласно стратегии В-2 при получении очередной копии ОР наравне с копиями из ВР используются и все ранее восстановленные копии массива данных.

В РСОД для обоих вариантов восстановления разрушенных данных можно использовать разные схемы восстановления, которые отличаются по времени, стоимости и вероятности успешного восстановления ОР. Например, если в узле j разрушены m копий массива данных (реплика и резерв), то в узле с ВР можно получить, а затем переслать в j-й узел y копий ($1 \leqslant y \leqslant m$), где остальные $(m - y)$ копий можно будет получить простым копированием в самом узле j. Различных вариантов восстановления множество, поэтому возникает задача выбора оптимальной схемы восстановления, т. е. нужно найти такое значение переменной y, которое обеспечило бы экстремум критерия оптимизации.

Рассмотрим данную задачу для случая использования восстановительной стратегии В-1. Будем использовать следующие критерии оптимизации: минимум средних затрат $Z(y)$, минимум среднего времени $E(y)$ и максимум вероятности $P(y)$ восстановления разрушенных данных. Предположим, что процесс получения y копий в узле с ВР и их передача по линиям связи производится последовательно.

В этом случае величина среднего времени $E(y)$ складывается из времени получения y копий в узле с ВР, времени t их передачи по каналам связи в узел j с разрушенными данными и времени копирования оставшихся $(m - y)$ копий в узле j.

Пусть T – среднее время создания одной копии в узле с ВР. При оценке среднего времени восстановления ОР необходимо рассмотреть два варианта (рис. 1):

а) $t \leqslant T$ – среднее время t передачи по каналам связи одной копии не превышает времени T ее создания в узле с ВР;

б) $t > T$ – время передачи t одной копии больше времени ее создания.

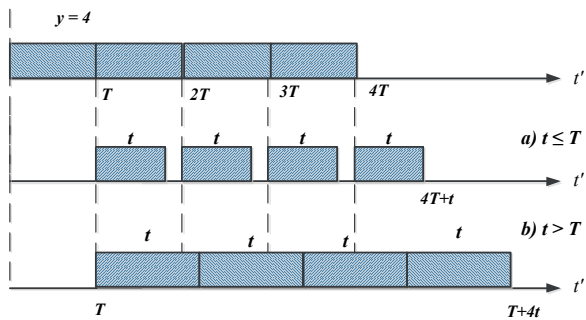

Рис. 1. Возможные соотношения между величинами T и t

С учетом этих двух вариантов среднее время $E(y)$ восстановления разрушенного ОР будет равно:

$$E(y) = \begin{cases} yT + t + (m - y)\,\tau \text{ при } t \leqslant T \\ yt + T + (m - y)\,\tau \text{ при } t > T. \end{cases} \quad (1)$$

Здесь τ – среднее время получения копии массива данных в узле с ОР.

В соответствии со сделанными предположениями вероятность $P(y)$ успешного восстановления ОР вычисляется по формуле

$$P(y) = \beta^m(\rho r\beta - 1), \quad (2)$$

здесь: ρ – вероятность успешного восстановления одной копии массива данных в узле с ВР; r – вероятность успешной передачи копии по каналам связи в узел сети с разрушенным ОР; β – вероятность успешного копирования массива в узле с разрушенным ОР.

Средние затраты на восстановление ОР будут равны:

$$Z(y) = y(D + d' - H) + Hm. \quad (3)$$

Здесь: D – стоимость восстановления одной копии массива данных в узле с ВР; d' - стоимость пересылки восстановленной копии массива в узел с разрушенным ОР; H – стоимость копирования массива в узле с разрушенным ОР.

Предположим, что $M \leqslant y \leqslant N (1 \leqslant M, N \leqslant m)$, и рассмотрим функцию $E(y)$.

Из (1) следует, что $E(y) < 0$ при $(t \leqslant T,\ T < \tau)$ и $(t > T,\ t < \tau)$. Тогда:

$$\begin{aligned} &\text{при } t \leqslant T \min E(y) = \begin{cases} m\tau + t + N(T - \tau) \text{ при } T < \tau \\ m\tau + t + M(T - \tau) \text{ при } T \geqslant \tau \end{cases} \\ &\text{при } t > T \min E(y) = \begin{cases} m\tau + T + N(t - \tau) \text{ при } t < \tau \\ m\tau + T + M(t - \tau) \text{ при } t \geqslant \tau. \end{cases} \end{aligned} \quad (4)$$

Для функций $P(y)$ и $Z(y)$ получим:

$$\max P(y) = \begin{cases} P(N) \text{ при } pr > \beta \\ P(M) \text{ при } pr \leqslant \beta \end{cases} \quad (5)$$

$$\max P(y) = \begin{cases} Z(N) \text{ при } pr > \beta \\ Z(M) \text{ при } pr \leqslant \beta. \end{cases} \quad (6)$$

Задача поиска оптимальной схемы восстановления разрушенных данных по критерию минимума стоимостных затрат на восстановление имеет формулировку:

$$Z(y) \to \min \tag{7}$$

при ограничениях:

$$P(y) \geqslant \bar{P}; \tag{8}$$

$$E(y) \geqslant \bar{E}; \tag{9}$$

$$t \in (1, 2, \ldots, m). \tag{10}$$

Можно показать, что ограничение (8) задачи эквивалентно одному из следующих двух ограничений[6]:

a) при $pr\beta^{-1} < 1$:

$$m \geqslant y \geqslant (\ln\bar{P} - m\ln\beta)(\ln p + \ln r - \ln\beta)^{-1} = A_1; \tag{11}$$

b) при $pr\beta^{-1} \geqslant 1$:

$$1 \leqslant y \leqslant (m\ln\beta - \ln\bar{P})(\ln\beta + \ln p - \ln r)^{-1} = A_2. \tag{12}$$

А ограничение (9) эквивалентно одному из следующих ограничений:

a) при $t \leqslant T$:

$$m \geqslant y \geqslant (t + m\tau - \bar{E})(t - T)^{-1} = B_1 \text{ при } T - \tau < 0; \tag{13}$$

$$1 \leqslant y \leqslant (\bar{E} - t - m\tau)(t - T)^{-1} = B_2 \text{ при } T - \tau > 0; \tag{14}$$

b) при $t > T$:

$$m \geqslant y \geqslant (T + m\tau - \bar{E})(\tau - t)^{-1} = B_1' \text{ при } t - \tau < 0; \tag{15}$$

$$1 \leqslant y \leqslant (\bar{E} - T - m\tau)(t - \tau)^{-1} = B_2' \text{ при } t - \tau > 0. \tag{16}$$

Величины определяют область допустимых значений для переменной y, в которой ищется решение сформулированной выше задачи (7) – (10). При этом для этих величин можно выделить 12 различных вариантов их значений, которые перечислены в таблице. При анализе данных вариантов, с учетом равенства (6), получаются все решения задачи, приведенные в таблице.

Таблица

№	Варианты значений ограничений задачи	Допустимые решения задачи	
		$D + d' < H$	$D + d' \geqslant H$
1	$[A_1] < [B_1] \leqslant y \leqslant m$	m	$[B_1]$
2	$[B_1] < [A_1] \leqslant y \leqslant m$	m	$[A_1]$
3	$[B_1] = [A_1] \leqslant y \leqslant m$	m	$[A_1] = [B_1]$
4	$1 \leqslant y \leqslant [A_2] < [B_2]$	$[A_2]$	1
5	$1 \leqslant y \leqslant [B_2] < [A_2]$	$[B_2]$	1
6	$1 \leqslant y \leqslant [B_2] < [A_2]$	$[A_2] < [B_2]$	1
7	$[A_1] \leqslant y \leqslant [B_2]$	$[B_2]$	$[A_1]$
8	$1 \leqslant y \leqslant [B_2]; [A_1] \leqslant y \leqslant m; [A_1] = [B_2]$	$[A_1] = [B_2]$	$[A_1] = [B_2]$
9	$1 \leqslant y \leqslant [B_2]; [A_1] \leqslant y \leqslant m; [B_2] = [A_2]$	решения нет	решения нет
10	$[B_1] \leqslant y \leqslant [A_2]$	$[A_2]$	$[B_1]$
11	$1 \leqslant y \leqslant [A_2]; [B_1] \leqslant y \leqslant m; [A_2] = [B_2]$	$[B_1] = [A_2]$	$[A_2] = [B_1]$
12	$1 \leqslant y \leqslant [A_2]; [B_1] \leqslant y \leqslant m; [A_2] < [B_2]$	решения нет	решения нет

Задача поиска оптимальной схемы восстановления разрушенных данных по критерию минимума среднего времени восстановления $E(y)$ имеет формулировку:

$$E(y) \to \min \tag{17}$$

при ограничениях:

$$Z(y) \leqslant \bar{Z}; \tag{18}$$

$$P(y) \geqslant \bar{P}; \tag{19}$$

$$y \in (1, 2, 3, ..., m). \tag{20}$$

Ранее было показано, что в зависимости от параметров системы ограничение (19) эквивалентно (11) или (12). Из (3) следует, что ограничение (18) эквивалентно одному из следующих неравенств:

$$m \geqslant y \geqslant (mH - \bar{Z})(H - D - d')^{-1} = C_1 \text{ при } D + d' < H; \tag{21}$$

$$1 \leqslant y \leqslant (\bar{Z} - mH)(D + d' - H)^{-1} = C_2 \text{ при } D + d' \geqslant H. \tag{22}$$

Аналогично задаче поиска оптимальной схемы восстановления по критерию минимума затрат, ограничения (11), (12) и (21), (22) определяют область значений переменной y, на которой ищется решение задачи.

В этом случае также возможны 12 ситуаций, анализ которых позволяет получить все решения задачи (17)–(20). Данные ситуации и соответствующие им решения показаны в таблице 1, в которой B_1, B_2 и $(D + d' < H)$, $(D + d' \geqslant H)$ необходимо заменить, соответственно, на C_1, C_2 и $(t \leqslant T < \tau$ или $\tau > t \geqslant T)$, $(t \leqslant T,\ T \geqslant \tau$ или $t > T,\ t \geqslant \tau)$.

Задача поиска схемы восстановления разрушенных данных, оптимальной по критерию максимума вероятности успешного восстановления, имеет формулировку:

$$P(y) \to \max \tag{23}$$

при ограничениях:

$$Z(y) \leqslant \bar{Z}; \tag{24}$$

$$E(y) \geqslant \bar{E}; \tag{25}$$

$$y \in (1, 2, 3, ..., m). \tag{26}$$

При анализе двух предыдущих задач было получено, что ограничения (24) и (25) в зависимости от параметров системы эквивалентны ограничению (21) или (22) и одному из ограничений (13)–(16) соответственно. Эти неравенства определяют область значений переменной y, на которой ищется решение задачи. Возможные при этом ситуации и соответствующие им решения задачи (23)–(26) приведены в таблице 1, в которой необходимо произвести следующие замены: A_1 и A_2 заменяются на C_1 и C_2, выражения $(D + d' < H)$ и $(D + d' \geqslant H)$ заменяются на $(pr\beta^{-1} > 1)$ и $(pr\beta^{-1} \leqslant 1)$ соответственно.

Рассмотрим второй тип восстановительного резервирования, когда для восстановления разрушенных данных в РСОД используются узлы с размещенным в них АМН, и получим условие, которое определяет целесообразность использования этого типа восстановительного резервирования с точки зрения времени восстановления разрушенных данных в РСОД.

Узлы с размещенным в них АМН обрабатывают только запросы на восстановление разрушенных данных в других узлах РСОД. В то время как узел с ОР в случае, если он используется в качестве

ВР (первый тип восстановительного резервирования), обрабатывает не только запросы на восстановление разрушенных данных, но также и информационные запросы и запросы на модификацию реплик.

Предположим, что РСОД построена на базе однородной сети, а для восстановления разрушенных данных можно использовать оба типа восстановительного резервирования: восстановление данных при помощи АМН и с помощью неразрушенного оперативного резерва ближайшего узла. Также предположим, что процесс обработки запросов на восстановление в узлах с АМН и ОР можно описать в терминах системы массового обслуживания вида M/M/1. Причем на вход данной системы поступает пуассоновский поток запросов, интенсивность которого для узла с ОР равна λ, а для узла с АМН равна μ. Время обслуживания запросов распределено по показательному закону. Будем считать, что время передачи сообщений по каналам связи (запроса на восстановление разрушенных данных и передачи восстановленных копий в узел с разрушенными данными) одинаково для обоих типов восстановительного резервирования. При сделанных предположениях среднее время T_A восстановления разрушенных данных при помощи АМН и аналогичное время $T_{\text{ОР}}$ при использовании неразрушенного ОР ближайшего узла РСОД будут равны[7]

$$T_A = T_\eta + E_A + \omega_A = T_\eta + E_A(1 - \mu E_A)^{-1} \tag{27}$$

$$T_{\text{ОР}} = T_\eta + E_{\text{ОР}} + \omega_{\text{ОР}} = T_\eta + E_{\text{ОР}}(1 - \lambda E_{\text{ОР}})^{-1}. \tag{28}$$

Здесь:

T_η – среднее время передачи информации по каналам связи;

E_A – среднее время обработки запроса на восстановление данных в узле АМН;

ω_A – среднее время ожидания запроса на восстановление в очереди на обработку в узле с АМН;

$E_{\text{ОР}}$ и $\omega_{\text{ОР}}$ – среднее время обработки и ожидания в очереди на обработку запроса на восстановление данных в узле с ОР;

μ – интенсивность запросов на восстановление данных, поступающих в узел сети с АМН;

λ – интенсивность всех запросов, поступающих в узел сети с неразрушенным ОР.

Очевидно, что с точки зрения величины среднего времени восстановления разрушенных данных использование для этой цели АМН нецелесообразно, если

$$T_{\text{ОР}} < T_A.$$

Это неравенство с учетом (27), (28) эквивалентно:

$$\lambda < \mu + (E_A - E_{\text{ОР}})(E_A E_{\text{ОР}})^{-1}. \quad (29)$$

Тогда можно сделать вывод о том, что если выполняется неравенство (29), то с точки зрения среднего времени восстановления разрушенных данных в узле РСОД целесообразно использовать в качестве восстановительного резерва неразрушенный резерв ближайшего узла вместо АМН.

Пример 1

Пусть РСОД функционирует на базе однородной полносвязной компьютерной сети, состоящей из N узлов, в которые поступают запросы с интенсивностью λ запросов в единицу времени для каждого узла.

В узлах РСОД размещен оперативный резерв по m копий массивов данных, созданный в соответствии со стратегией 1 резервирования. Вероятность разрушения одной копии массива данных при ее использовании для обработки запроса равна q.

Определим целесообразность создания в одном из узлов РСОД архива магнитных носителей объемом m копий с точки зрения величины среднего времени восстановления разрушенных данных.

Так как объем (m) оперативного резерва, созданного в узлах РСОД, совпадает с объемом АМН, а компьютерная сеть однородна, то получим, что $E_A = E_{\text{ОР}}$, а условие (29) примет вид

$$\lambda < \mu. \quad (30)$$

Так как в соответствии с нашими условиями оперативный резерв в узлах РСОД создан согласно стратегии 1 резервирования, то очевидно, что интенсивность μ запросов на восстановление разрушенного ОР, поступающих в узел с АМН, будет равна

$$\mu = \sum_{i=1} \lambda q^m = N\lambda q^m.$$

Тогда условие (30) будет эквивалентно следующему неравенству

$$1 < Nq^{\eta 2}. \quad (31)$$

Рассмотрим теперь два частных случая.

1) Пусть $m = 3$ и $N = 4$.

В этом случае использование АМН в РСОД нецелесообразно с точки зрения величины среднего времени восстановления разру-

шенных данных при $q > 0{,}69$. Однако на практике в реальных системах обработки данных $q < 1/2$. Из этого следует, что при заданных параметрах использование АМН будет целесообразно.

2) Пусть $m = 3$.

Для данного случая определим число N узлов РСОД с размещенными в них репликами и резервом данных, при котором использование только одного узла с АМН для их восстановления будет нецелесообразно в том смысле, что время ожидания запроса на восстановление данных в очереди на обслуживание в узле с АМН будет больше, чем в случае, если для восстановления данных будем использовать узел с неразрушенным ОР вместо узла с АМН.

Из выражения (31) следует, что при $m = 3$ получим $N > q^{-3}$. Тогда при $q < 1/2$ создание и размещение в компьютерной сети только одного АМН будет нецелесообразно для $N \geqslant 8$, а при $q = 0{,}1$ для $N > 1000$.

Теперь рассмотрим ситуацию, когда в РСОД используется несколько АМН, размещенных в узлах компьютерной сети. Предположим, что все возникающие запросы на восстановление разрушенных данных распределяются между всеми узлами с АМН равномерно. Пусть K – это число узлов с АМН.

Определим, какое число K узлов с АМН должно быть в РСОД, чтобы среднее время восстановления разрушенных данных с помощью узла с АМН было меньше, чем с помощью ближайшего узла с неразрушенным ОР. Из (30) следует, что для этого необходимо не менее μ/λ узлов с АМН. Используя неравенство (31), получим, что при условии

$$K \geqslant Nq^{\eta 1} \qquad (32)$$

обработка запросов на восстановление разрушенных данных с помощью узлов с АМН будет производиться за меньшее время, чем с помощью узлов с неразрушенным оперативным резервом.

Пример 2

Предположим, что $N = 30$, $m = 2$, $q = 0{,}2$. При данных параметрах из (32) следует, что в РСОД необходимо разместить АМН объемом в 2 копии в $K = 2$ узлам РСОД.

В статье рассмотрен метод повышения надежности распределенных систем обработки данных с использованием восстановительного резервирования. Сформулирована задача выбора оптимальной схемы восстановления разрушенных реплик и резерва массива данных с помощью двух восстановительных стратегий. Проведен анализ областей допустимых решений задачи. Выполнен

анализ использования двух вариантов восстановительного резерва: неразрушенный резерв ближайшего узла системы и архив магнитных носителей. Полученные результаты позволяют до 15% снизить ущерб от потери информации в распределенных системах. Результаты работы целесообразно использовать при определении оптимальной схемы восстановления разрушенной информации при проектировании архитектуры распределенных систем обработки данных, функционирующих на базе ненадежных компьютерных сетей.

Примечания

1 *Казарин О.В.* Безопасность программного обеспечения компьютерных систем. М.: МГУЛ, 2003.

2 *Таненбаум Э., ван Стеен М.* Распределенные системы. Принципы и парадигмы. СПб.: Питер, 2003.

3 *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.

4 *Микрин Е.А., Сомов С.К.* Анализ эффективности стратегий восстановления информации в распределенных системах обработки данных // Информационные технологии и вычислительные системы. 2016. № 3. С. 5–19.

5 *Кульба В.В., Сомов С.К., Шелков А.Б.* Резервирование данных в сетях ЭВМ. Казань: Изд-во Казанского ун-та, 1987.

6 *Сомов С.К.* Резервирование программных модулей и информационных массивов в сетях ЭВМ: Дис. ... канд. техн. наук. М.: ИПУ РАН, 1983.

7 *Клейнрок Л.* Вычислительные системы с очередями. М.: Мир, 1979.