

В. Н. Ушаков

РЕАЛИЗАЦИЯ МЕТОДОВ ОБЕСПЕЧЕНИЯ НАДЕЖНОСТИ В БАЗАХ ДАННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

Описываются методы, которые были разработаны и применены при практическом решении задач обеспечения надежности баз данных информационных систем: метода параметрического резервирования протоколов транзакций, комбинированного пермутационного метода поиска данных, метода оптимизации алгоритмов установки контрольных точек процессов. Приводятся результаты экспериментального исследования применения методов на этапах разработки, внедрения и эксплуатации автоматизированного комплекса на базе ЭВМ СМ-1600, предназначенного для задач, решаемых в процессе управления газовым хозяйством и локализации аварийных ситуаций.

§ 1. ПРИНЦИПЫ МОДЕЛИРОВАНИЯ НАДЕЖНОСТИ БАЗ ДАННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

Современные автоматизированные информационные системы и системы управления, как правило, являются сложными программно-техническими комплексами, эффективность функционирования которых возможно только при обеспечении достаточной их надежности. Требуемая эффективность программно-технических комплексов информационных систем может быть достигнута путем разработки и внедрения специальных методов обеспечения надежности систем. При этом возникает необходимость, в частности, решения следующих проблем:

сохранения больших информационных массивов в условиях периодической их корректировки и расширения;

внедрения специальных методов поиска информации в больших массивах с различной структурой представления данных;

обеспечения достаточно высокой степени экономической эффективности системы;

адекватного представления объектов в массивах данных информационной системы

В Москве в промышленную эксплуатацию внедряется вторая очередь организационно-технологической автоматизированной системы управления «Мосгаз», в состав которой входят подсистемы «Локализация аварийных ситуаций» и «Управление эксплуатацией газового хозяйства». Создание комплексов программно-технического обеспечения и технологических инструкций эксплуатации подсистем вызвали необходимость разработки и внедрения ряда методов обеспечения надежности и экспериментального исследования эффективности этих методов

При разработке методов учитывалось, что природа информационной системы как коммуникационной системы, функционирующей обязательно непосредственно или косвенно с участием человека, всегда двойственна. С одной стороны, это алгоритмическая система, созданная для решения определенных задач и выполняющая множество заданных алгоритмов. С другой стороны, в большинстве случаев, это диалоговая система, в которой реализуются определенные принципы организации человеко-машинного комплекса с некоторой, иногда значительной, долей интуитивности в выборе событий.

Информационная система может рассматриваться как продукт целенаправленной научно-технической и организационно-технологической деятельности человека или определенной группы людей. Поэтому при организации процессов разработки, эксплуатации и анализа системы

в полной мере могут применяться, с одной стороны, структура понятий и положения традиционной теории надежности, обычно применяемые при анализе характеристик технических устройств, а с другой, категории организации процессов эвристического приятия решений.

Кроме этого, при моделировании системы базы данных всегда следует учитывать, что всякая вновь создаваемая система обязательно в цикле своей жизни проходит, по крайней мере, два этапа развития: этап отладки и этап эксплуатации. С учетом этого должно проводиться моделирование информационной системы, при выборе методов анализа обязательно должны рассматриваться условия стационарности параметров и характеристик исследуемых компонент и системы в целом.

В [1; 2] определяются состав и общие правила задания требований к надежности технического изделия. Номенклатура показателей надежности включает показатели безотказности, долговечности, ремонтпригодности, сохраняемости и комплексные показатели. Каждый показатель надежности характеризуется специальными более подробными показателями: показателями сохранения эффективности, вероятностью безотказной работы, коэффициентами готовности изделия к эксплуатации и т. д.

При выборе модели надежности программного обеспечения системы можно воспользоваться классификацией моделей надежности, представленной в [3]. Существует два типа моделей. Модели возрастания надежности применяются в процессе отладки изделия, модели надежности, наоборот, применяются после завершения процесса отладки. Первые позволяют моделировать тестирование, отказы и процессы устранения отказов при отладке системы, предсказывать, когда среднее время наработки на отказ окажется достаточно большим, чтобы можно было начать эксплуатацию программы. Вторые применяются, когда при тестировании системы не обнаружены ошибки, процесс отладки завершен и операционное качество программного продукта оказывается на допустимом для эксплуатации уровне.

Надежность является статистическим показателем отказов системы, которые возникают из-за различных дефектов. Модели надежности позволяют предсказывать среднюю величину ожидаемого потока отказов. Однако традиционная теория надежности, которая была разработана в свое время в связи с обеспечением надежности проектирования физических объектов, не может считаться удовлетворительной для всех случаев исследования программного обеспечения. Она может приме-

няться только в случаях, когда выполняются множества аналогичных операционных предположений.

Модели возрастания надежности и модели надежности отличаются качественно. Так как первые используются в процессе отладки, когда наблюдаются отказы, тестирование дает непосредственную статистическую обратную связь показателей надежности с моделируемой производительностью системы. Средняя величина потока отказов, рассчитываемая с помощью этой модели, оказывается более грубой, чем в действительности, потому что расчет ведется на основании данных короткого периода, например, периода, когда произведено определенное число запусков программ. Более того, поскольку модель возрастания надежности разрабатывается лишь с целью указать момент, когда операционное качество программного обеспечения оказывается на допустимом уровне, то точность предсказания может быть значительно ниже, чем точность в модели надежности.

Предсказания теории надежности, с другой стороны, слабее связаны со статистическими характеристиками. Поскольку отказов может не наблюдаться, нет никакой возможности контроля предсказываемого потока отказов. Более того, среднее время наработки на отказ может быть много больше времени наблюдения. Между тем, в системах обеспечения безопасности требования к времени наработки на отказ могут быть чрезвычайно высокими. Для таких систем рассчитываемое среднее время наработки на отказ должно вычисляться точно, потому что система должна обеспечивать выполнение контрактных условий или сохранение жизни людей.

В § 2 описывается метод параметрического резервирования протоколов транзакций баз данных, обеспечивающий контроль показателей надежности баз данных с учетом случайного характера возникновения аппаратных отказов. Представлены результаты внедрения этого метода в программное обеспечение действующей системы.

В § 3 описывается комбинированный пермутационный метод поиска данных. Приводятся результаты экспериментов, позволяющие оптимизировать процесс поиска информации в массивах текстовых данных.

В § 4 приводится описание метода оптимизации алгоритмов установки контрольных точек процессов. Излагаются результаты практического внедрения метода в программном обеспечении комплекса на базе ЭВМ СМ-1600, позволяющего учитывать случайные остановки процессов и обеспечивающего повышение экономической эффективности базы данных коллективного пользования.

Метод параметрического резервирования протоколов транзакций баз данных и метод оптимизации алгоритмов установки контрольных точек, с учетом сказанного выше, относятся к методологии моделей возрастания надежности и могут применяться в задачах обеспечения надежности информационных систем на этапах их разработки, отладки и эксплуатации. Эти методы разработаны с учетом лишь «алгоритмической» части требований и не позволяют учитывать основанные на интуитивных принципах эвристические факторы в процессах жизненного цикла системы. Тем не менее, как показывают результаты их практического внедрения, представленные методы оказались весьма полезными при организации процессов разработки, внедрения и эксплуатации автоматизированного комплекса на базе ЭВМ СМ-1600, предназначенного для решения важных задач экономии материальных ресурсов и предотвращения аварийных ситуаций в Москве.

Комбинированный пермутационный метод поиска данных, напротив, относится к методологии моделей надежности и позволяет прогнозировать эффективность эксплуатации информационной системы в условиях не-

достаточности или даже отсутствия статистических данных о результатах тестирования автоматизированного комплекса.

§ 2. МЕТОД ПАРАМЕТРИЧЕСКОГО РЕЗЕРВИРОВАНИЯ ПРОТОКОЛОВ ТРАНСАКЦИЙ БАЗ ДАННЫХ

При эксплуатации систем управления базами данных (СУБД) возникает задача обеспечения требуемого уровня сохранности системы и устранения последствий сбоев, обусловленных неисправностями аппаратных средств. В случае разрушения файлов ядра базы данных (БД), методика дублирования информации с помощью стандартных утилит СУБД не всегда должным образом обеспечивает сохранение информации, введенной в БД после очередного сеанса резервирования данных. Для обеспечения высокой сохраняемости БД при возникновении аппаратных сбоев и повышения эффективности работы систем БД должны применяться специальные методы резервирования информации протоколов транзакций.

Предлагаемый метод параметрического резервирования протоколов транзакций предполагает реализацию соответствующего алгоритма и следует специальной методике расчета степени дублирования данных, которые подробно описаны в [4].

Алгоритм параметрического резервирования протоколов транзакций (ПРПТ) предназначен для разработки прикладных программ систем БД. В нем реализован известный метод дублирования данных, но этот метод применяется в данном случае с учетом требований эксплуатации БД коллективного пользования. Основное отличие представленного ниже алгоритма ПРПТ от стандартных алгоритмов ведения журнала СУБД заключается в способе записи протоколов транзакций БД: стандартный протокол СУБД (например, СУБД ADABAS-RSX) обычно записывается средствами СУБД в журнал ядра БД и заносится в архив при архивации БД; протокол предлагаемого алгоритма ПРПТ ведется на уровне подпрограмм прикладного программного обеспечения и записывается на резервный магнитный диск (МД) в виде набора файлов.

Алгоритм ПРПТ обеспечивает сохраняемость набора информации при разрушении БД из-за аппаратных сбоев и допускает отбор записей протокола по наборам параметров: датам корректировки записей БД, именам пользователей, специальным идентификаторам записей, значениям полей записей, что существенно расширяет эксплуатационные возможности СУБД.

Распределение вероятности отказов БД по времени эксплуатации описывается распределением Пуассона (при этом предполагается, что учитываются только случайные факторы отказов, а сами события отказов не зависят одно от другого):

$$P(T) = \frac{(W \times T)^k}{k!} \times \exp(-W \times T),$$

где $P(T)$ — вероятность отказа за время T ; W — параметр потока отказов; k — количество одновременных отказов.

В качестве критерия эффективности средств резервирования можно предложить X — максимально допустимую долю потерь данных протоколов. Величина X определяется допустимым отношением затрат на дублирование S и непроизводительных затрат из-за потерь при сбоях Z к общим затратам на производимые операции J :

$$X = (S + Z) / J.$$

Другой смысл величины X заключается в том, что для периода, равного общему времени работы системы, величина X составляет вероятность отказа за этот период [5]. Отсюда получаем формулу критерия необходимой степени дублирования данных:

$$X = (S + Z) / J > \frac{(W \times T)^k}{k!} \times \exp(-W \times T) \quad (1)$$

По этой формуле можно определить, достаточно ли часто дублируется информация протоколов БД.

Описанный метод параметрического резервирования протоколов транзакций реализован в программном обеспечении системы БД коллективного пользования на СМ-1600.

В табл. 1 представлены показатели избыточности

Таблица 1

Показатели избыточности хранения информации при параметрическом резервировании протоколов в подсистеме «Сифон» по данным за 11 месяцев 1993 г.

Номера массивов	Объем архива протоколов базы данных в 512-байтных блоках по макету 1	Объем архива протоколов базы данных в 512-байтных блоках по макету 10	Коэффициент избыточности хранения информации
1, 2	4834	1952	1,18
3, 4	4067	2322	1,18
5, 6	3972	2104	1,18
7, 8	3322	2060	1,18
9, 10, 11	4920	2925	1,18
Итого:	21 115	11363	1,18

хранения информации при параметрическом резервировании протоколов транзакций БД подсистемы «Сифон». Подсистема входит в состав комплекса автоматизированной организационно-технологической системы управления газовым хозяйством Москвы, реализованной на базе ЭВМ СМ-1600 и персональных компьютерах IBM AT.

Приведенные сведения иллюстрируют относительное увеличение пространства на МД, занятого сохраняемыми файлами протоколов из-за обязательного дублирования ключевых (дескрипторных) реквизитов при выгрузке записей БД.

В табл. 2 приводятся сведения о количестве аппарат-

Таблица 2

Количество аппаратных отказов с потерей протоколов транзакций базы данных подсистемы «Сифон» без резервирования за 11 месяцев 1992 г.

Номера массивов	Общее количество записей в массивах	Количество сеансов работы с базой данных	Общее количество аппаратных отказов	Количество «фатальных» отказов с потерями протоколов
1, 2	18 891	65	10	3
3, 4	16 196	47	6	2
5, 6	15 985	54	10	4
7, 8	13 288	38	8	2
9, 10, 11	19 831	75	25	16
Итого:	84 191	279	59	27

ных отказов с потерей протоколов транзакций БД подсистемы «Сифон».

Первая очередь подсистемы была разработана к началу 1992 г. и не предусматривала включения в состав комплекса технических и программных средств автоматизированного резервирования информации протоколов БД.

В табл. 2 сведения основаны на статистических данных, собранных за 11 месяцев 1992 г. в журнале эксплуатации подсистемы. Общее количество отказов аппаратных средств составило $N_0=59$. Количество «фатальных» отказов аппаратных средств, приведших к потерям информации и к необходимости повторения ручного ввода информации в БД, составило $N=27$ при общем числе сеансов работы пользователей с базой данных $M=279$.

Во второй очереди комплекса подсистемы «Сифон», которая эксплуатируется с января 1993 г., предусмотрена возможность автоматизированного параметрического резервирования протоколов транзакций по представленному выше методу ПРПТ.

В табл. 3 приводятся сведения о количестве аппарат-

Таблица 3

Количество аппаратных отказов с потерей протоколов транзакций базы данных подсистемы «Сифон» с резервированием за 11 месяцев 1993 года

Номера массивов	Общее количество записей в массивах	Количество сеансов работы с базой данных	Общее количество аппаратных отказов	Количество «фатальных» отказов с потерями протоколов
1, 2	27242	58	13	2
3, 4	27682	57	19	1
5, 6	27263	67	13	—
7, 8	25347	73	17	2
9, 10, 11	32719	67	18	4
Итого:	140 243	322	80	9

ных отказов с потерей протоколов транзакций подсистемы «Сифон» второй очереди.

Представленные в табл. 3 сведения основаны на статистических данных, собранных за 11 месяцев 1993 г. в журнале эксплуатации подсистемы. Общее количество отказов аппаратных средств составило $N_0=80$. Количество «фатальных» отказов аппаратных средств, приведших к потерям информации и к необходимости повторения ручного ввода информации в БД, составило $N=9$ при общем числе сеансов работы пользователей с базой данных $M=322$.

Сравнение представленных экспериментальных данных и результатов расчетов, полученных по формуле (1), демонстрируют хорошую, хотя и не абсолютную, эффективность применяемого во второй очереди подсистемы «Сифон» метода ПРПТ.

Алгоритм резервирования протоколов транзакций, реализованный по методу ПРПТ, позволил примерно в 3,5 раза уменьшить долю непроизводительных издержек, связанных с имевшими место аппаратными отказами комплекса. Расхождение экспериментального результата с расчетным составило около 40%, что находится в пределах ошибки вычислений при расчетах, составляющей в конкретном рассматриваемом случае также 40%.

§ 3. КОМБИНИРОВАННЫЙ ПЕРМУТАЦИОННЫЙ МЕТОД ПОИСКА ДАННЫХ

При разработке программ поиска данных в файлах с текстовой информацией иногда требуется определить достаточно эффективный алгоритм нахождения записи файла в условиях, когда существует неопределенность структурной организации записи или когда заданный критерий поиска по каким-либо причинам не вполне определен.

При автоматизированном поиске информации в БД часто пользователь-оператор бывает вынужден формировать запрос на основании неполных или нечетко сформулированных наборов входной информации. Например, при обращении клиентов к пользователю-оператору ВС клиент может не знать точного адреса или фамилии искомого лица или знать их лишь приблизительно. При передаче оперативной информации о критериях поиска данных в системах радиосвязи или телефонной связи могут оказываться сбои, ухудшающие качество приема сведений. Для предотвращения снижения надежности ВС в этих и подобных случаях разработчики программного обеспечения при подготовке массивов данных могут воспользоваться комбинированным пермутационным методом поиска данных [6].

Этот метод основан на трех методах повышения надежности программ ВС: методе фонетического и семантического расширения полей дескрипторов (ФРД), методе морфологического сокращения полей дескрипторов (МСД) и методе поиска данных с ограниченной пермутацией и последовательным символьным перебором полей дескрипторов (ППД).

При фонетическом расширении полей дескрипторов (ФРД) происходит расширение массива поиска по наборам гласных и согласных звуков — в массив вводятся дополнительные записи. На основе проведенного фонетического анализа, в состав поискового массива включаются дополнительные фрагменты текстовой информации, в которых учитываются возможные искажения звучания исходных фрагментов данных.

Расширение поискового массива по возможным семантическим вариантам написания строк записей текстового массива производится путем перебора возможных способов письменного обозначения объектов поиска.

Метод морфологического сокращения полей дескрипторов (МСД) применяется для повышения вероятности безотказного поиска данных. Сокращение значений полей может производиться на основании данных морфологического анализа слов в полях данных массива поиска.

Метод поиска данных с ограниченной пермутацией и последовательным символьным перебором полей дескрипторов (ППД) предполагает последовательный перебор вариантов поиска по символам дескрипторного поля с перестановкой слов, включенных в поле дескриптора.

Пример. Пусть в массиве поиска имеется фрагмент:

ул. Туристская дом 16 участок 2

Предположим, поступил запрос о данных по улице имени Вилиса Лациса, новое название которой *ул. Туристская* (Москва). Оператор ЭВМ может не знать нового названия улицы, поэтому в поисковый массив целесообразно заранее внести как старое, так и новое название улицы. Кроме того, ограниченная пермутация строк по номерам домов и участков позволяет осуществлять поиск также и по номерам домов и участков:

<i>ул. Туристская</i>	<i>дом 16</i>	<i>участок 2</i>
<i>ул. Вилиса Лациса</i>	<i>дом 16</i>	<i>участок 2</i>
<i>дом 16</i>	<i>участок 2</i>	<i>ул. Туристская</i>
<i>участок 2</i>		<i>ул. Туристская дом 16</i>

В последних двух строках примера поиск может осуществляться, например, по критериям поиска: «16» или «участок 2». При этом ограниченная пермутация фрагментов позволяет осуществлять последовательный символьный перебор фрагментов строк по ограниченному числу символов (например, по 1—16 символам), без перебора всей строки, что сокращает число операций, выполняемых программой, то есть ускоряет поиск. Морфологическое сокращение (усечение) критерия поиска, например, с фрагмента *Вилиса Лациса* до фрагмента *Лацис* также значительно ускорит поиск. В [6] показано, что наилучшие результаты поиска получаются при длине фрагмента критерия поиска в 4—5 символов.

Описанные выше методы ФРД, МСД и ППД рекомендуются применять совместно, именно в этом случае обеспечивается качественно новый подход к организации поиска данных и реализуется комбинированный пермутационный метод.

Рассмотрим относительное изменение средней по времени вероятности безотказного поиска данных (ВБПД) при использовании комбинированного пермутационного метода при выборе подхода к определению критериев поиска и формированию массивов. При строгом рассмотрении эта вероятность определяется многими факторами, в том числе, субъективными, в частности, эргономическими факторами и условиями оперативной ситуации вокруг пользователя ВС.

Смысл этой вероятности заключается в том, что она является количественной мерой надежности ВС при поиске данных, поскольку она, в простейшем случае рассмотрения, определяет показатель безотказности ВС, который, в свою очередь, является показателем надежности работы человеко-машинного комплекса. ВБПД можно представить в виде функции:

$$P = F(N, M, m, L, n, S, r), \quad (2)$$

где N — число записей в массиве поиска данных; M — число букв во фрагменте критерия поиска; m — число букв во фрагменте критерия поиска, вызывающих сомнения в правильности написания слова; L — длина строки записи в поисковом массиве; n — доля записей в поисковом массиве данных с буквами, вызывающими сомнения в написании слов; S — среднее число возможных вариантов написания букв, вызывающих сомнения в правильности написания слова; r — среднее число возможных правильных вариантов написаний букв, вызывающих сомнения в правильности написания слова.

С одной стороны, величина P приблизительно равна отношению количества успешных сеансов поиска к общему количеству попыток поиска, поскольку это отношение при усреднении по времени работы определяется отношением времени безотказной работы к общему времени работы системы линейным образом [5].

С другой стороны, величина P может быть определена через среднюю вероятность отказа G :

$$P = 1 - G = F(N, M, m, L, n, S, r) = p.$$

Предположим, что средняя вероятность отказа при поиске, $G = g$, определена для случая, когда среднее число возможных правильных вариантов написаний букв $r = 0$. Тогда величина P при поиске без применения метода ФРД может быть выражена в виде

$$P_1 = p \times (1 - (g \times n \times (S^m - r)) / S^m). \quad (3)$$

При поиске в массиве всегда имеется хотя бы один правильный фрагмент критерия поиска, т. е. в обычном массиве всегда $r = 1$. В случае использования метода ФРД и создания дополнительного массива с наиболее вероятными ошибочными значениями дескрипторов выполняется соотношение:

$$r \approx S^m.$$

Поэтому относительное изменение средней ВБПД при использовании метода ФРД, по сравнению с обычным способом поиска, можно оценить по формуле

$$PS = (p-p \times (1 - (g \times n \times (S^m - r)) / S^m)) / g = n \times (S^m - r) / S^m \quad (4)$$

В табл. 4 приводятся результаты расчета изменений

Таблица 4

Относительное увеличение ВБПД при использовании методов ФРД и ППД

Параметр	Значения параметров расчетов		
n	0,04	0,10	0,50
S	2	2	2
m	1	1	2
r	1	1	1
PS	0,02	0,05	0,375

средней ВБПД по формуле (4). Из приведенных данных можно заключить: относительная средняя вероятность безотказного поиска данных значительно увеличивается при совместном применении методов ФРД и ППД. Применение указанных методов к адресным массивам Москвы может обеспечить существенное значение величины PS . Метод ФРД применялся лишь к 4% записей, поскольку лишь небольшое количество строк при фонетическом анализе адресного массива вызвали сомнения в возможности правильного определения фрагмента критерия поиска. При этом по адресным массивам Москвы реальное значение PS составило 2%. Тем не менее, следует ожидать, что применение метода ФРД при поиске в массивах данных, в которых используются иностранные слова и фамилии, значительное число иноязычных слов, будет весьма эффективным. Как показывают расчеты по формуле (4), в случае фонетического расширения 50% записей массива с добавлением в среднем по две записи, величина PS может увеличиться до 37,5%. Ошибка расчета определяется ошибкой определения отношения среднего времени безотказной работы к общему времени работы системы и, в случае применения метода ФРД к 4% записей, не превышает величины в 2,5%.

§ 4. МЕТОД ОПТИМИЗАЦИИ АЛГОРИТМОВ УСТАНОВКИ КОНТРОЛЬНЫХ ТОЧЕК ПРОЦЕССОВ

При длительном выполнении процессов на ЭВМ возникает задача сохранить промежуточные результаты расчетов на магнитных дисках или в памяти ЭВМ, чтобы предотвратить потери рабочего времени из-за остановок или сбойных ситуаций. Решение этой технической задачи связано с решением научной проблемы: при выборе интервалов установки контрольных точек сохранения промежуточных результатов расчетов необходимо, во-первых, определить критерий оптимальности выбора временных параметров процессов, во-вторых, обеспечить выполнение требований оптимальности временных параметров процессов по выбранному критерию [7].

Установку контрольных точек (КТ) будем рассматривать в простейшем случае конвейерных вычислений с

единственным процессом в операционной системе. Каждая КТ процесса должна включать описания основных расчетных параметров процессов: промежуточные суммы, показатели счетных циклов и т. д. Критерием оценки эффективности метода является минимизация непродолжительных вычислительных затрат.

Введем основные обозначения:

t — затраты времени на установку одной КТ;

T — период установки КТ;

T_{Π} — продолжительность одного процесса;

$T_{\text{Б}}$ — среднее время безостановочной (безотказной) работы вычислительной системы;

T_0 — последнее время выполнения части процесса после последней КТ в процессе;

N — число запусков процессов за время $T_{\text{Б}}$;

$T_{\text{н}}$ — время непроизводительных затрат за время $T_{\text{Б}}$;

W — средний поток отказов (остановок);

k — коэффициент непроизводительных затрат;

n — число контрольных точек, устанавливаемых за один процесс.

Продолжительность одного процесса:

$$T_{\Pi} = T_0 + n \times T.$$

Кроме того,

$$T_{\text{Б}} = \frac{1}{W} = T_{\Pi} \times N \times (1 - W \times \exp(-W \times T_{\Pi} \times N)).$$

При $n \gg 1$ и $N \gg 1$ имеем $\frac{T_0}{T_{\text{Б}}} \ll 1$. Поэтому

$$T_{\text{Б}} = \frac{1}{W} \approx T_{\Pi} \times N.$$

Будем считать, что за один процесс в среднем происходит не более одного отказа. Тогда получаем

$$T_{\text{н}} = t \times n \times N + W \times T \times \exp(-W \times T) \times T_0.$$

Поэтому коэффициент непроизводительных затрат:

$$k = \frac{T_{\text{н}}}{T_{\text{Б}}} \approx \frac{t}{T} \times \frac{n}{\left(n + \frac{T_0}{T}\right)} + W \times T \times \exp(-W \times T) \times \frac{T \times W \times \left(n + \frac{T_0}{T}\right)}{\left(\frac{n \times T}{T_0} + 1\right)}.$$

С учетом того, что $n \gg 1$, $\frac{T_0}{T} \approx \frac{1}{2}$, получаем:

$$k \approx \frac{t}{T} + \frac{1}{2} \times (W \times T)^2 \times \exp(-W \times T). \quad (5)$$

Второе слагаемое в формуле (5) учитывает случайный характер событий прерывания процессов, независимо от причины прерываний. Выведенной формулой можно успешно пользоваться для оценочных расчетов, проводимых при $n \gg 1$ и $N \gg 1$.

Описанная методика установки контрольных точек по рассчитанным выше временным параметрам была реализована в программном обеспечении подсистемы «Управление эксплуатацией газового хозяйства». В табл. 5 приводятся сведения о результатах эксплуатации программ подсистемы с установкой контрольных точек процессов. За период эксплуатации программ с 1 февраля по 31 декабря 1993 г. произошло 35 отказов и преднамеренных остановок процессов. При этом экономия рабочего времени комплекса ЭВМ составила около 22 часов. (Ошибка расчета не превышает 20%).

Таблица 5

Сведения о результатах эксплуатации программ с установкой контрольных точек за 11 месяцев в период с 1 февраля по 31 декабря 1993 года

Наименование задачи	Общее количество отказов и остановок задачи за период	Среднее время выполнения процесса (час)	Общая экономия машинного времени (час.)
«ANO»	4	16	12
«BEV»	7	1,5	6
«REC»	24	0,5	4
Итого:	35	—	22

ЗАКЛЮЧЕНИЕ

Представленный метод параметрического резервирования протоколов транзакций баз данных (ПРПТ) является основой достаточно простого и эффективного алгоритма дублирования информации протоколов транзакций баз данных с учетом случайного характера возникновения сбойных ситуаций в средствах технического и программного обеспечения. Предложенный критерий достаточности степени дублирования информации при резервировании протоколов БД позволяет контролировать непроизводительные вычислительные затраты. Экспериментальная проверка расчетов показала, что предложенный метод ПРПТ обеспечивает поддержку показателей сохранности БД в заданных пределах непроизводительных затрат из-за аппаратных отказов с точностью не хуже 40%.

Применение описанного комбинированного пермутационного метода поиска данных (КПМ) приводит к значительному увеличению относительной средней вероятности безотказного поиска данных в текстовых мас-

сивах. Метод КПМ может быть особенно эффективен в системах поиска с массивами данных, в которых используются иностранные слова и фамилии, массивы иноязычных слов.

Предложенный метод оптимизации алгоритмов установки контрольных точек процессов позволяет контролировать величину непроизводительных затрат машинного времени с учетом случайного характера прерывания процессов из-за отказов ЭВМ или преднамеренных остановок процессов оператором. При реализации метода в программных средствах с длительным выполнением вычислительных процессов достигается заметная экономия машинного времени. Значительный экономический эффект от внедрения метода получается, если обеспечен интенсивный многопользовательский режим работы в базах данных коллективного доступа.

СПИСОК ЛИТЕРАТУРЫ

1. Правила безопасности в газовом хозяйстве.— М.: Недра, 1991.— с. 118—127.
2. Надежность в технике. Состав и общие правила задания требований к надежности РД 50-650-87. Госстандарт, М., 1988.
3. Hamlet D. «Are We Testing for True Reliability? // IEEE Software.— 1992.— July.— P. 21—27.
4. Ушаков В. Н. Метод параметрического резервирования протоколов транзакций баз данных.— М.: 1994.— Деп. в ВИНТИ 24.01.94, N 204-B94.
5. Атовмян И. О. и др. Надежность автоматизированных систем управления.— М.: Высшая школа. 1979.— с. 16.
6. Ушаков В. Н. Методы обеспечения надежности поиска информации в системах баз данных.— М.: 1993.— Деп. в ВИНТИ 15.02.93, N 360-B93.
7. Ушаков В. Н. Метод оптимизации алгоритмов установки контрольных точек процессов.— М.: 1994.— Деп. в ВИНТИ 24.01.94, N 205-B94.

Материал поступил в редакцию 18.05.94.

НАШИ АВТОРЫ

МАТЕЕВ Николай — кандидат философских наук, старший научный сотрудник Института социологии Болгарской Академии наук

МАТЕЕВ Пламен — кандидат математических наук, старший научный сотрудник Института математики Болгарской Академии наук

ШЕКЕРДЖИЙСКА Йорданка — научный сотрудник Центра наукознания Болгарской Академии наук

ВОЙСКУНСКИЙ Владимир Григорьевич — кандидат технических наук, старший научный сотрудник Российской национальной библиотеки, С-Петербург

ШМАИН Илья Ханаанович (выпускник механико-математического факультета МГУ, бывший сотрудник ВИНТИ) — священник русской православной церкви в Сен-Женевьев де Буа, Франция

УШАКОВ Валерий Николаевич — ведущий инженер МП «Мосинформсистема», аспирант ВИНТИ, Москва

Редактор **Т. Н. Лаппалайнен**

Технический редактор **Л. В. Кутакова**

ЛР № 040228 от 22.01.92 Подписано в печать 12.07.94 Сдано в набор 03.06.94
Бум. тип. № 2 Формат бумаги 84×108^{1/16} Литературная гарнитура Высокая печать
Усл. печ. л. 3,36 Уч.-изд. л. 5,70 Тираж 784 экз Заказ 2155 Цена 550 р.
Адрес редакции: 125219, Москва, ул. Усневича, 20а, Тел. 152-66-71
Производственно-издательский комбинат ВИНТИ,
140010, г. Люберцы, Московской обл., Октябрьский пр-т, 403