

МЕТОДЫ ПОВЫШЕНИЯ ОТКАЗОУСТОЙЧИВОСТИ ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМ

Введение

Основной задачей телекоммуникационных систем является расширение сферы предоставляемых услуг и повышение их качества. Очевидно, что успешное решение этой задачи возможно только при высокой экономической эффективности работы сети, а именно – в случае значительной надёжности и помехоустойчивости передачи наряду с низкой стоимостью передаваемого информационного бита. Всё это делает актуальным развитие таких направлений в технике телекоммуникаций, как разработка совершенных устройств и методов уплотнения и подходов к повышению достоверности и надёжности передачи.

Эффективность управления телекоммуникационными инфраструктурами на базе современных информационных, сетевых и компьютерных технологий требует создания мультисервисных сетей телекоммуникаций общего пользования и телематических служб, образованных терминалами многофункционального и интеллектуального типа с повышенной пропускной способностью для удовлетворения постоянно растущих потребностей общества в телекоммуникации. Решение этой проблемы неразрывно связано с повышением уровня технического обеспечения телекоммуникационных систем (ТС), реализованного многофункционального терминального комплекса для интеграции процессов обслуживания и управления передачей информации.

При этом задача повышения уровня технического обеспечения систем телекоммуникаций (СТ) решается путем создания многофункциональных и интеллектуальных терминальных комплексов с внедрением инфокоммуникационных технологий, которые относятся к наиболее актуальным задачам. Такая масштабная информатизация в первую очередь связана с повышением качества функционирования ТС, главным звеном которых являются многофункциональные терминальные комплексы на основе функционально-блочных систем (модемы, шлюзы, интерфейсы, концентраторы, мосты, терминальные мультиплексоры и элементы управления), представляющие собой распределённые системы управления на базе TMN (Telecommunication Management Network).

Исследуемые *эффективность и помехоустойчивость* системы являются комплексными характеристиками *качества функционирования ТС*, образованных терминалами многофункционального и интеллектуального типа, отражают способность системы обеспечивать заданные технические, экономические и эксплуатационные характеристики при выполнении возложенных на нее абонентских, информационных и сетевых функций в рамках системы обмена речевой, видео и неречевой информации.

Среди ученых стран СНГ и других зарубежных стран, занимающихся проблемами *повышения эффективности и помехоустойчивости* СТ и их технического обеспечения, можно выделить работы В.А. Котельникова, Г.П. Захарова, В.П. Бакалова, Н.Б. Зелигера, В.П. Шувалова, Л.Н. Финка, В.Н. Гордиенко, Ф.Г. Мамедова, О.С. Чугреева, В.О. Шварцмана, В.А. Игнатова, А.Н. Гасанова, Н.В. Захарченко, П. Беккера, И.А. Ушакова, Э.А. Якубайтиса, L. Kleinrock, K. Shannon, O. Serlin, D. Ting, M. Schwartz и др.

Работы указанных авторов подготовили базу для развития нового научного направления в создании систем передачи и приема цифровой информации, связанного с разработкой методов расчета показателей качества функционирования технических средств в сетях телекоммуникации. Однако в этих работах недостаточно внимания уделено созданию многофункциональных ТК для передачи и приёма речевой и неречевой информации при их совместном обслуживании. Кроме того, эти решения не охватывают всей проблемы повышения эффективности и помехоустойчивости ТК СТ.

В направлении повышения эффективности элементов и устройств вычислительной техники и систем управления ТС важные результаты получены в работах Б.П.Терентьева, В.В.Губарева, С.И.Сахарчука, В.Н.Четверикова, Т.М.Алиева, В.С.Шибанова Н.Г.Фарзана, В.В.Лебедянцева, В.П.Дьяконова., К.М.Имамвердиева, Р.Т.Гумбатова, J.Grabner, J.Wieselmann и др. Однако применение этих результатов для улучшения качества функционирования ТК с внедрением современных технологий возможно при существенных ограничениях на эффективность использования канальных и терминальных ресурсов.

Значительный вклад в фундаментальные исследования, направленные на создание новых ЭВМ и их использование в пионерских системах управления технологическими процессами, энергетическими и другими объектами, в том числе военного назначения, для систем автоматизации научного эксперимента, был сделан академиком В.М. Глушковым, основоположником информатики в Украине, и многими ведущими учеными основанного им Института кибернетики НАН Украины.

Важнейшей проблемой является повышение отказоустойчивости телекоммуникационных систем (ТС). В соответствии с [1] отказоустойчивость основана на реализации в полном или усеченном виде цепочки действий:

- прогнозирование (fault forecasting) возможности появления (проявления) дефекта и возникновения отказа вследствие этого дефекта;
- предупреждение (fault prevention) появления (проявления) дефекта и возникновения отказа;
- обнаружение (fault detection) появления (проявления) дефекта, ошибки вычислений, отказа;
- идентификация (fault diagnosis) причины, вида и места дефекта (отказа);
- парирование (fault tolerance) последствия дефекта и возникновения отказа.

Последнее действие может включать: реконфигурацию (fault removal) структуры (архитектуры) путем исключения отказавшего компонента из конфигурации, замены работоспособным и восстановление вычислительного процесса (fault recovery). Механизмы отказоустойчивости должны быть инвариантны по отношению к типам дефектов, вызывающих отказы.

Одним из важных свойств телекоммуникационных систем и систем управления критическими объектами является отказоустойчивость. Классически под отказоустойчивостью вычислительной (телекоммуникационной) системы понимается способность устранения последствий неисправностей элементов и программного обеспечения за счет использования аппаратной, информационной и алгоритмической избыточности [2].

Повышение отказоустойчивости предполагает подавление влияния отказов и сбоев на работу ТС. Оно осуществляется с помощью средств контроля и коррекции ошибок, а также средств автоматического восстановления вычислительного процесса после проявления неисправности, включая аппаратную, программную и временную избыточность, на основе которых реализуются различные варианты отказоустойчивых архитектур. Повышение отказоустойчивости – это способ предотвращения отказа системы. Поэтому разработка новых методов повышения отказоустойчивости ТС является актуальной задачей.

Основные методы обеспечения отказоустойчивости ТС

Наряду с термином «отказоустойчивая система» в эквивалентном смысле встречаются термины «системы высокой готовности», «устойчивые и эластичные к сбоям и отказам системы», «системы непрерывной и постоянной готовности». Однако термины трактуются по-разному отдельными разработчиками, поставщиками и потребителями. Рассмотрим трактовки некоторых терминов.

Высокая Готовность (High Availability). Все системы высокой готовности обеспечивают устойчивость к отказам и сбоям благодаря применению принципа избыточности. Конфигурация системы обеспечивает ее быстрое восстановление после обнаружения неисправности, для чего используются избыточные аппаратные и программные средства. Для современных

систем высокой готовности характерно использование технологии “горячей” замены отказавшего узла.

Эластичность к сбоям (Fault Resiliency). Ряд специалистов делят весь диапазон систем высокой готовности на две части, при этом в верхней части оказываются системы, эластичные к сбоям. Эластичность к сбоям определяет более короткое время восстановления, которое позволяет системе быстро «откатиться» назад при обнаружении неисправности.

Устойчивость к сбоям (Fault Tolerance). Устойчивые к сбоям системы (“отказоустойчивые системы”) имеют в своем составе избыточную аппаратуру для всех функциональных блоков, включая процессоры, источники питания, подсистемы ввода/вывода и подсистемы дисковой памяти. Если соответствующий функциональный блок неправильно работает, всегда имеется горячий резерв, и неисправность в любом блоке не может вывести систему из строя. В ряде отказоустойчивых систем избыточные аппаратные средства можно использовать для распараллеливания обычных работ.

Непрерывная готовность (Continuous Availability). Лучшими среди отказоустойчивых систем являются системы, обеспечивающие непрерывную готовность. Система с непрерывной готовностью, если она работает корректно, позволяет ликвидировать как плановые, так и неплановые простои и проводить модернизацию (upgrade) и обслуживание системы в оперативном режиме (режиме on-line). Разработка подобной системы охватывает как аппаратные средства, так и программное обеспечение. Очень важным дополнительным требованием к таким системам является сохранение уровня производительности в случае отказа какого-либо компонента.

Следует отметить, что понятие «непрерывная готовность» все-таки подразумевает возможность сбоя в работе системы. Однако если типичное значение коэффициента готовности для обычной вычислительной (телекоммуникационной системы) системы составляет около 99 %, то для систем непрерывной готовности этот параметр устанавливается на уровне не менее 99.999 %.

Именно системы непрерывной готовности обеспечивают наивысший возможный уровень гарантии *работоспособности системы*. Помимо того, что непрерывная готовность является проверенной технологией, она представляет собой простое, простое в использовании, прозрачное для приложений и базирующееся на стандартах решение.

Постоянная готовность (Permanent Availability). Это скорее теоретическое понятие. Коэффициент готовности для такого рода систем принимается равным 100 %, т.е. предполагается отсутствие сбоев в работе, что практически недостижимо.

Выбор того или иного уровня готовности определяется теми последствиями, которые будут иметь для конкретного объекта отказы управляющей системы.

Фирмы-изготовители предлагают специальные средства для построения *отказоустойчивого нижнего уровня ТС*. Типичным является применение специальных модулей для построения дублированных управляющих систем (УС) с «горячим» резервированием. Структура подобной системы приведена на рис. 1.

Она содержит два идентичных контроллера (основной и резервный), арбитр и схему коммутации выходных сигналов. Входные сигналы от объекта поступают в оба контроллера.

Широко применяются также троированные структуры. На рис. 2 приведена схема, иллюстрирующая работу троированной системы. Как и в предыдущем примере, входные сигналы и другая информация поступают одновременно в три идентичных между собой контроллера, в которых синхронно выполняется одна и та же программа. После очередного цикла получения результатов для управления объектом они от каждого контроллера поступают на мажоритарную схему «2 из 3», формирующую интегральный результат. *Отказоустойчивость обеспечивается на всех уровнях:* от сети управления до блоков питания плат ввода-вывода. Для достижения максимальной надежности системы обеспечивается резервирование следующих компонентов и функций: ЛВС, контроллеров, системы SCADA, станции архивной регистрации и сервера отчетов, операторских рабочих станций, сетевых серверов и концентраторов.

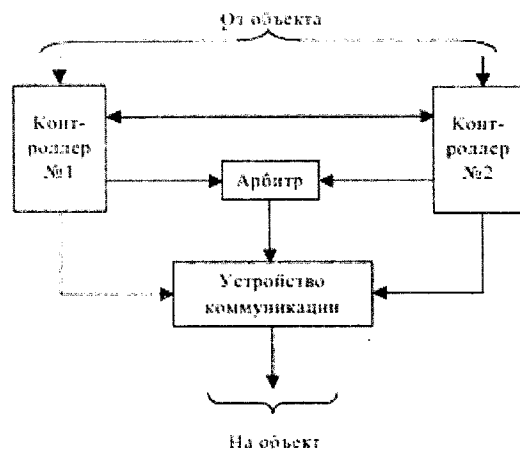


Рис. 1. Структура дублированной УС

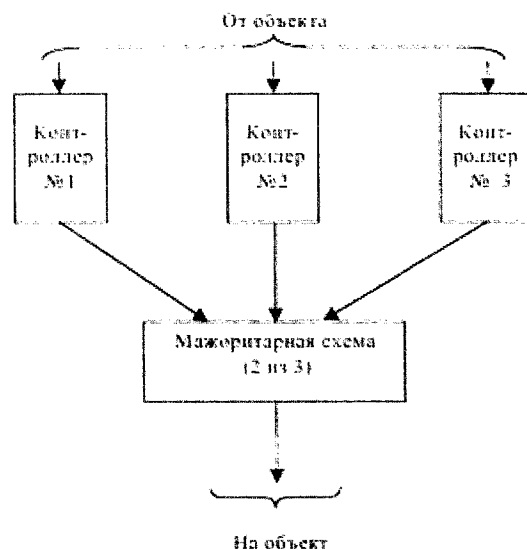


Рис. 2. Структура троированной системы

Отказоустойчивость ЛВС осуществляется протоколом, обеспечивающим гарантированную передачу данных по сети, и мощными свойствами самовосстановления. Резервируемая сеть конфигурируется в виде двойного кольца, обеспечивающего реконфигурацию при отказе любого сегмента магистрали. При реконфигурации изменяется маршрутизация технологических данных, что обеспечивает бесперебойную работу системы. Оба кольца сети постоянно находятся в активном состоянии, причем одно из них работает как сеть данных, а другое – как диагностическая сеть. В редких случаях отказов первой сети вторая захватывает поток данных, гарантируя их целостность.

Системные сообщения передаются по обеим сетям для снижения статистической частоты ошибок. Оперативное обнаружение отказов и сбор данных о производительности коммуникационных каналов системы выполняются индивидуально для каждого устройства. При обнаружении отказа устройство, в котором возникла проблема, автоматически отключается от остальной системы до устранения неисправности.

Каждое устройство в сети обладает несколькими резервированными функциями, что обеспечивает максимальную надежность системы в целом.

Например, все устройства системы управления имеют дублированную схему подключения к каждой из сетей данных, в результате чего достигается четырехкратное резервирование. Контроллер «Овации» можно сконфигурировать с несколькими резервированными функциями, включая дублирование материнских плат процессора, источников питания и ка-

налов связи. Уровень резервирования контроллера выбирается в соответствии с техническими требованиями заказчика. При дублировании процессоров обеспечивается основное и резервное управление с полностью безударным автоматическим переключением между процессорами при сбое или отказе. Автоматическое включение резерва может инициироваться широким набором событий, включая: отказ управляющего процессора; отказ сетевого адаптера; отказ интерфейса ввода – вывода; пропадание питания в управляющем процессоре; перезапуск управляющего процессора.

Как только управление переходит к резервному процессору, неисправный процессор можно отключить, отремонтировать и подключить обратно без оказания воздействия на выполнение алгоритмов управления. После перезапуска восстановленный процессор обнаруживает, что его партнер находится в режиме управления и принимает на себя резервную роль. Управляющий процессор обнаруживает наличие резервного процессора и настраивается на работу с резервом.

Для построения отказоустойчивого верхнего уровня ТС могут использоваться отказоустойчивые серверы. Так, в общепромышленном секторе нашли применение отказоустойчивые серверы Stratus Technologies [3].

В основе технологии Stratus Continuous Processing лежат три фактора:

1. Аппаратная отказоустойчивость с резервированием всех компонентов. При этом используется жёсткая (потактовая) синхронизация процессов и данных в дублирующих друг друга компонентах. Все компоненты архитектуры Continuous Processing работают параллельно и согласованно для того, чтобы не просто минимизировать время незапланированного простоя, а предотвратить его возникновение. Таким образом, в случае отказа одного из них потери данных и остановки сервера не происходит.

2. Программное обеспечение повышенной устойчивости. В сервере используются программные средства для предотвращения случайных сбоев в аппаратуре.

3. Архитектура активного сервиса. Подразумевается возможность горячей замены любого компонента (включая процессоры) без остановки приложений. После замены компоненты автоматически включаются в работу.

Конструкция заменяемых блоков исключает ошибку при снятии/установке, поэтому замена может производиться пользователем самостоятельно. Сервер автоматически извещает центр поддержки о происшествии или неисправности, после чего, при необходимости, нужный блок доставляется для замены.

Критерии отказоустойчивости ТС

При оценке эффективности функционирования ТС возникает необходимость выделения подсистем и других элементов для описания поведения их в системе. Каждая из подсистем ТС исследуется в соответствии со своими частными критериями. Показано [4], что если отдельные подсистемы функционируют относительно независимо, то оптимизация по частным критериям подсистем обеспечивает оптимальное функционирование подсистемы в целом, когда критерий эффективности системы является монотонной функцией эффективности функционирования каждой из подсистем. В общем случае, когда имеет место сложная корреляция процессов функционирования различных элементов системы (а именно так обстоит дело на практике) это может не соблюдаться.

Естественно, что эффективность ТС, как сложной системы, не может быть оценена по одному критерию. Значение каждого из показателей качества ТС в той или иной мере характеризует эффективность функционирования ТС, т.е. степень соответствия его назначению – максимально быстро удовлетворять требования по вводу, обработке и выводу информации. Причем, между ними может существовать взаимосвязь.

Таким образом, критерий эффективности ТС является векторным. На практике его применять неудобно. Общий подход к решению проблемы выбора критерия эффективности предложен в [4]. В соответствии с этим подходом на практике можно использовать следующие варианты:

1. Из множества показателей следует выбрать один, наиболее важный.

2. Можно попытаться уменьшить размерность критерия, введя взвешенную сумму различных показателей, при этом, однако, выбор численных значений весовых коэффициентов остается трудной задачей.

3. Можно попытаться отыскать некоторый более общий показатель, поглощающий все остальные частные показатели системы. Очень часто такой показатель является критерием эффективности надсистемы.

В соответствии с третьим вариантом можно предложить интегральный подход, когда эффективность какой-либо подсистемы оценивается через степень влияния ее характеристик на показатели эффективности всей системы, в которую входит данная подсистема. Такой подход оправдан тем, что, в конечном итоге, для системы важны показатели ее подсистемы, рассматриваемой как «черный ящик». Например, для системы управления технологическим объектом важными являются интегральные характеристики всего ТС в целом, т.е. его внешние характеристики, а множество внутренних характеристик его подсистем и компонентов является малозначимым.

Для подсистем ТС, выполняющих расчетные задачи моделирования в реальном или ускоренном масштабах времени, главным качеством является достижение необходимой производительности, при которой цикл моделирования выполняется за заданный промежуток времени. Невыполнение цикла моделирования за заданный промежуток времени фактически является *отказом* подсистемы. И наоборот, гарантированное выполнение цикла моделирования за заданный промежуток времени является одним из свойств ее *отказоустойчивости*. Подобные подсистемы ТС, в частности, применяются в системах поддержки принятия решения, в моделирующих вычислительных комплексах, предназначенных для отработки систем управления космическими аппаратами и бортовых систем управления [5].

Производительность применяется для оценки качества вычислительных систем, одной из разновидностей которых является ТС. К настоящему времени сформулированы несколько различных показателей производительности вычислительной системы (ВС) [6 – 9]. Такими показателями могут быть, например:

- производительность, рассчитанная как среднее число задач, которые могут быть выполнены в единицу времени;
- относительная пропускная способность – усредненное отношение числа обслуженных запросов к числу поступивших в единицу времени;
- коэффициент использования (отношение среднего времени работы ВС к общему времени функционирования с учетом простоев).

Еще большее многообразие понятий наблюдается в определениях параметров производительности ВС при выполнении конкретных прикладных программ, например [68-70], от средних значений времен выполнения отдельных операций, времен отклика ВС на различные запросы до времен выполнения "смесей" команд и/или процедур, времени выполнения классических численных и/или "эталонных" задач.

Зачастую упомянутые показатели рассматриваются как критерии эффективности ВС общего применения для различных приложений.

Такой показатель, как номинальная производительность, характеризует "потенциальную" возможность ВС, т.е. максимально возможную физическую скорость обработки данных. Реальная или эффективная производительность зависит от множества факторов, как заложенных при проектировании ВС (соотношение характеристик быстродействия ее компонентов, дисциплина их обслуживания), так и являющихся характеристиками операционных систем и прикладных программных процессов конкретных систем контроля и управления (характеристики потоков данных, степень "накладных расходов" при работе подсистемы ввода-вывода операционной системы). Эффективная производительность ВС равна произведению номинальной производительности и коэффициента (вероятности) загрузки ВС.

При известной номинальной производительности коэффициент (вероятность) загрузки можно использовать в качестве показателя эффективной производительности ВС. Под коэф-

фициентом загрузки имеется в виду отношение полезного времени работы системы к общему времени функционирования системы с учетом простоев. Коэффициент загрузки характеризует реальную пропускную способность системы. Коэффициент загрузки широко применяется как критерий качества при исследовании ВС общего назначения.

Увеличение коэффициента загрузки системы при фиксированной номинальной производительности увеличивает пропускную способность системы. Для моделирующей подсистемы это выражается в увеличении вероятности гарантированного выполнения цикла моделирования за заданный промежуток времени, т.е. в увеличении *отказоустойчивости*.

В связи с этим возникает задача определения коэффициента загрузки подсистем ТС.

ТС функционируют в реальном масштабе времени. Это означает, что времена отклика ТС $T_{\text{юнк}}$ по каждой i -й функции управления не должны превышать предельных интервалов времени $T_{\text{инр}}$, определяемых спецификой управляемого объекта. Если $T_{\text{юнк}} > T_{\text{инр}}$, то в этом случае ухудшается качество управления, что может привести к нарушению работы управляемого объекта.

Времена отклика ТС являются суммарным временем нескольких составляющих: времени преобразования аналогового сигнала в цифровой код, первичной обработки, времени вычислительных процедур соответствующей задачи, передачи результатов обработки в модули управления исполнительными механизмами, в рабочие станции для визуализации, архивирования и пр. Следовательно, и для каждой временной составляющей существует предельное ограничение.

Поэтому, одним из главных критериев отказоустойчивости подсистем ТС является способность выполнять i -ю функцию за время, не превышающее заданное директивное время $T_{\text{дир}}$, с требуемой вероятностью D_i .

Например, для коммуникационной подсистемы ТС это означает, что заявки по передаче сообщений от задачи одного абонента линии вычислительной сети (ЛВС) к задаче другого абонента должны обслуживаться за какое-то ограниченное время с требуемой вероятностью.

Для прикладной программы (приложения), состоящей из набора задач и выполняющей одну из функций ИУС, это означает выполнение ее за время, не превышающее $T_{\text{дир}}$, заданное проектировщиком ИУС, с требуемой вероятностью D . В частности, для подсистем ТС, выполняющих расчетные задачи моделирования в реальном или ускоренном масштабах времени, это означает, что цикл моделирования должен выполняться за время, не большее заданного директивного времени $T_{\text{дир}}$ с требуемой вероятностью D .

В работе [10] для управляющих ВС реального времени данный критерий трактуется как «надежное выполнение набора задач управления».

В понятии «надежного выполнения задач» отказ управляющего ТС трактуется не только как потеря его работоспособности, но и как невыполнение требуемых задач за время $T_{\text{дир}}$. Такой отказ может иметь место даже при безотказном функционировании аппаратуры и программного обеспечения ТС, отвечающего принципам «высокая готовность» (High Availability) и «непрерывная готовность» (Continuous Availability). Причинами такого отказа в некорректно спроектированных ТС могут быть неучтенные проектировщиком информационных управляющих систем (ИУС) внезапные изменения параметров объекта управления из-за случайных динамических изменений условий его функционирования или параметров окружающей среды, случайные изменения объемов данных и объемов вычислений по программам задач, задержки выполнения приложений вследствие некоторых типов прерываний, конфликты процессов при обращении к общим ресурсам и пр.

Время выполнения задач может также существенно увеличиваться и превышать директивное время из-за случайных неисправностей (сбоев или отказов) компонентов ТС.

Вследствие случайного характера указанных явлений и эффектов, времена выполнения задач и их фрагментов (программных модулей), а также наборов задач в целом необходимо рассматривать, в общем случае, как случайные величины, а процесс выполнения наборов задач в управляющих ТС – как случайный процесс.

В связи с этим возникают *вопросы оценки вероятности успешного завершения набора задач* за заданное директивное время в ТС с заданной производительностью вычислительных ресурсов, прогнозирования успешного завершения набора задач непосредственно при их выполнении в ТС, управления вычислительными процессами для предотвращения угрозы отказа.

Поэтому возникает задача прогнозирования времени выполнения набора задач на вычислительных ресурсах ТС.

В некоторых работах (например, [11-13]) это время рассматривается как совокупность длительностей между событиями, связанными с выполнением заданной программы, либо как некоторая относительная величина, характеризующая изменение предполагаемого времени выполнения программы относительно некоторого эталона. Известны также многочисленные отечественные и зарубежные разработки и исследования по прогнозированию или даже точному расчету времени выполнения наборов задач (в том числе – взаимосвязанных) в параллельных ВС – в предположении, что время решения каждой задачи заранее известно, т.е. является детерминированным (а неслучайным, как в рассматриваемом нами подходе); к таким работам можно отнести, в первую очередь, фундаментальные и комплексные исследования, проведенные в [14]).

Применительно к управляющим ВС, включая управляющие ТС, проблема оценок такого рода получила название математического прогнозирования времени выполнения сложных наборов задач (программных комплексов) и нашла решение в [10, 15-18].

Принципиально, что эти наборы рассмотрены в наиболее трудоемких версиях – с произвольными информационными и логическими связями между задачами – и потому названы комплексами взаимосвязанных работ (КВР) – задач и/или их параллельно-последовательных фрагментов (подзадач, процессов), а в общем случае – программных модулей. Во многих публикациях (например, в [19]) аналогичные структуры именуются как “графы задач”. В предлагаемом подходе КВР следует рассматривать как графовую модель рабочей нагрузки.

В работах [10, 15 – 18] под математическим прогнозированием времени выполнения конкретного, заданного пользователем КВР понимается определение в статике (т.е. до реализации задач в ТС) стохастических оценок времени T реализации КВР (среднего значения, дисперсии, функции $F(t)$ распределения времени T) и определение вероятности D завершения КВР за время, не превышающее заданное $T_{дир}$, в ТС с заданной конфигурацией и производительностью его вычислительных ресурсов.

Другой задачей, требующей решения, является задача определения параметров производительности ТС реального времени, в которой набор из S задач выполнялся бы за время, не большее директивного времени $T_{дир}$ с заданной вероятностью D .

Составной частью исследования по обеспечению надежного выполнения взаимосвязанных задач в вычислительных системах реального времени является проблема диспетчеризации (планирования, расписания) наборов задач (программных модулей) по вычислительным ресурсам ВС.

В настоящее время существует большое количество публикаций, посвященных решению этой проблемы. Традиционный интерес к данной проблеме и ее актуальность обусловлены в основном двумя факторами:

- возможностью получения существенного эффекта – заметного уменьшения общего времени выполнения заданного набора задач – при незначительных издержках на разработку и функционирование конкретного алгоритма диспетчеризации;
- непрекращающимся стремлением к “синтезу” оптимальных дисциплин диспетчеризации по различным критериям.

В данной работе на основе методологии [16 – 18] необходимо решить задачу формальной оценки и выбора дисциплин динамической диспетчеризации работ КВР по вычислителям ТС при случайных временах выполнения работ КВР – с целью минимизации времени выполнения каждого заданного КВР, что должно повысить отказоустойчивость ТС.

В связи с этим возникают вопросы оценки вероятности успешного завершения набора задач за заданное директивное время в ТС с заданной производительностью вычислительных ресурсов, прогнозирования успешного завершения набора задач непосредственно при их выполнении в ТС, управления вычислительными процессами для предотвращения угрозы отказа.

Поэтому возникает задача прогнозирования времени выполнения набора задач на вычислительных ресурсах ТС.

В некоторых работах (например, [11-13]) это время рассматривается как совокупность длительностей между событиями, связанными с выполнением заданной программы, либо как некоторая относительная величина, характеризующая изменение предполагаемого времени выполнения программы относительно некоторого эталона. Известны также многочисленные отечественные и зарубежные разработки и исследования по прогнозированию или даже точному расчету времени выполнения наборов задач (в том числе – взаимосвязанных) в параллельных ВС – в предположении, что время решения каждой задачи заранее известно, т.е. является детерминированным (а неслучайным, как в рассматриваемом нами подходе); к таким работам можно отнести, в первую очередь, фундаментальные и комплексные исследования, проведенные в [14]).

Применительно к управляющим ВС, включая управляющие ТС, проблема оценок такого рода получила название математического прогнозирования времени выполнения сложных наборов задач (программных комплексов) и нашла решение в [10, 15-18].

Принципиально, что эти наборы рассмотрены в наиболее трудоемких версиях – с произвольными информационными и логическими связями между задачами – и потому названы комплексами взаимосвязанных работ (КВР) – задач и/или их параллельно-последовательных фрагментов (подзадач, процессов), а в общем случае – программных модулей. Во многих публикациях (например, в [19]) аналогичные структуры именуются как “графы задач”. В предлагаемом подходе КВР следует рассматривать как графовую модель рабочей нагрузки.

В работах [10, 15 – 18] под математическим прогнозированием времени выполнения конкретного, заданного пользователем КВР понимается определение в статике (т.е. до реализации задач в ТС) стохастических оценок времени T реализации КВР (среднего значения, дисперсии, функции $F(t)$ распределения времени T) и определение вероятности D завершения КВР за время, не превышающее заданное $T_{дир}$, в ТС с заданной конфигурацией и производительностью его вычислительных ресурсов.

Другой задачей, требующей решения, является задача определения параметров производительности ТС реального времени, в которой набор из S задач выполнялся бы за время, не большее директивного времени $T_{дир}$ с заданной вероятностью D .

Составной частью исследования по обеспечению надежного выполнения взаимосвязанных задач в вычислительных системах реального времени является проблема диспетчеризации (планирования, расписания) наборов задач (программных модулей) по вычислительным ресурсам ВС.

В настоящее время существует большое количество публикаций, посвященных решению этой проблемы. Традиционный интерес к данной проблеме и ее актуальность обусловлены в основном двумя факторами:

- возможностью получения существенного эффекта – заметного уменьшения общего времени выполнения заданного набора задач – при незначительных издержках на разработку и функционирование конкретного алгоритма диспетчеризации;
- непрекращающимся стремлением к “синтезу” оптимальных дисциплин диспетчеризации по различным критериям.

В данной работе на основе методологии [16 – 18] необходимо решить задачу формальной оценки и выбора дисциплин динамической диспетчеризации работ КВР по вычислителям ТС при случайных временах выполнения работ КВР – с целью минимизации времени выполнения каждого заданного КВР, что должно повысить отказоустойчивость ТС.

Особенности данного исследования, отличающие ее от известных работ по аналогичной проблематике, заключаются в следующем:

- дисциплины и диспетчеризации исследуются на основе методологии [17 – 19] математического прогнозирования времени выполнения сложных наборов задач с произвольными информационными и логическими взаимозависимостями между задачами, называемыми комплексами взаимосвязанных работ;

- времена выполнения работ могут быть различны и рассматриваются как случайные величины;

- рассматривается выполнение КВР в неоднородных (несимметричных) вычислительных подсистемах ТС с параллельными вычислительными ресурсами; в этом случае одни и те же работы могут выполняться на вычислителях разных типов, но имеющих различную производительность, а следовательно, – с различными временами выполнения одной и той же работы.

Важными критериями отказоустойчивости ТС также являются [20]:

- высокая надежность его функционирования;

- контроль и управление всеми параметрами критической системы в реальном масштабе времени;

- обеспечение диагностирования и восстановления работоспособности всех подсистем за допустимое разработчиком время;

- обеспечение требуемой глубины диагностирования ТС или его компонентов в реальном масштабе времени за минимальное число тестовых проб или наблюдений.

Поэтому необходимо решить задачу разработки и исследования принципов организации процессов диагностирования неисправностей вычислительных средств ТС и использования дополнительных средств асинхронного резервирования процессов, которые должны обеспечивать не только *диагностирование* и локализацию неисправностей вычислительных ресурсов параллельных ТС и корректное восстановление процессов, но и включать математические методы оценки влияния случайных неисправностей на время выполнения сложных наборов задач.

Необходимо также исследовать и разработать методы, приводящие к существенному уменьшению времени тестирования и диагностирования ТС в процессе эксплуатации на основе создания общей модели организации и проведения диагностического эксперимента, включающего условные и безусловные алгоритмы поиска дефектов с использованием IEEE стандартов тестопригодного проектирования.

Следующим важным критерием отказоустойчивости подсистем ТС, функционирующих в реальном времени, является готовность подсистемы. В [1] готовность (availability) определена как свойство доступности ресурсов системы для предоставления требуемых услуг. Например, в отношении ЛВС под этим термином следует понимать отношение суммарного времени пребывания системы в состояниях, когда ЛВС не работает из-за отсутствия заявок либо, когда идет передача сообщения, и при этом нет других заявок, ожидающих обслуживания, к общему времени функционирования системы [21].

Под готовностью подсистемы Q будет пониматься величина, равная отношению времени работы данной подсистемы, когда в ТС нет заявок, ожидающих обслуживания данной подсистемой.

Заключение

ТС является сложной отказоустойчивой системой, состоящей из подсистем, которые могут включать большое количество взаимодействующих элементов. Для его исследования будут применяться методы аналитического и имитационного моделирования.

Динамику выполнения в ТС программных приложений, реализующих новые функции ИУС, можно представить как набор взаимосвязанных последовательно-параллельных программных модулей (задач, работ) или комплекс взаимосвязанных работ (КВР).