

ОБЗОРЫ

УДК 621.327

С.П. Сущенко

ИССЛЕДОВАНИЯ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ И СЕТЕЙ В ТОМСКОМ ГОСУДАРСТВЕННОМ УНИВЕРСИТЕТЕ

Разработан комплекс моделей сетевых структур и сформулированы принципы построения магистральных трактов передачи данных. Созданы модели многоуровневой памяти и исследовано влияние ассоциативности кэш-памяти на быстродействие вычислительной системы.

Совокупность функциональных моделей вычислительных сетей и систем включает модели следующих сетевых структур и компонентов вычислителей:

1. Замкнутые модели межузловое соединения, позволяющие учитывать влияние искажений потока протокольных блоков данных в прямом и обратном каналах связи на операционные характеристики звена передачи данных и выполнять оптимизацию по критерию пропускной способности межузловых соединений значений длины кадра и ширины окна.

2. Открытые модели звена передачи данных в виде двухзвенного сетевого фрагмента, позволяющие исследовать влияние фактора блокировок буферной памяти транзитного узла коммутации и протокольных параметров на производительность стартопного и конвейерного протоколов управления звеном передачи данных.

3. Открытые модели многозвенного тракта передачи данных, позволяющие анализировать влияние качества каналов связи и емкости буферных накопителей транзитных узлов коммутации на операционные показатели протокола сетевого уровня.

4. Детерминированные конвейерные модели многозвенного тракта, позволяющие исследовать процесс передачи потока мультимедийных сообщений по неоднородным многозвенным виртуальным соединениям на транспортном уровне управления сетью. На основе моделей данного процесса, отличающихся учетом конвейерного эффекта, возможно решение задачи оптимальной фрагментации абонентских сообщений на пакеты данных. В рамках этих моделей реально разработать методов определения длины кадра и ширины окна линейного уровня неоднородной сети передачи данных, совместно учитывающих требования держателей средств связи к пропускной способности межузловых соединений и требования пользователей к задержке абонентских сообщений.

5. Стохастические модели конвейерных механизмов информационного переноса, позволяющие исследовать влияние длительности тайм-аута ожидания квантации транспортного уровня на сквозную задержку сообщений в виртуальном канале и синтезировать длительность сквозного тайм-аута, обеспечивающие заданный уровень вероятности повторной передачи данных.

6. Дискретные модели процесса соперничества для случайного множественного доступа совокупности информационных источников к разделяемой среде передачи данных.

7. Модели многоуровневой памяти вычислительных систем, определяющие влияние параметров ассоциативности и глубины неблокируемости кэш-памяти на операционные показатели быстродействия вычислителя.

1. Замкнутые модели протоколов канального уровня

Построены замкнутые модели нормальных и асинхронных процедур управления звеном передачи данных для группового и селективного режимов защиты от ошибок [1–2], обладающие преемственностью по отношению к стартопному протоколу. Для нор-

мальных процедур обмена показано, что пропускная способность имеет мультипликативную форму зависимости от коэффициента информационной скорости прикладных данных в детерминированном межузловом соединении и коэффициентов, определяющих влияние фактора искажений информационных протокольных блоков данных в прямом и фактора подтверждений в обратном каналах связи. Данный операционный показатель для асинхронных протоколов имеет более сложную функциональную связь с факторами искажений в прямом и обратном каналах.

При анализе моделей асинхронных управляющих процедур установлено, что для однородного дуплексного канала связи возможности широко используемого режима группового отказа, соответствующие неограниченной ширине окна, в режиме селективного отказа достигаются уже при ширине окна, равной двум.

Для нормальных и асинхронных процедур управления звеном передачи данных получены аналитические оценки оптимальных по пропускной способности размеров кадра однородной и неоднородной сети и значений размера окна различных управляющих процедур от объема накладных расходов и качества канала связи. Показана инвариантность оценок оптимальной длины кадра к режиму защиты от ошибок нормальных процедур обмена. Исследовано качество полученных оценок. Предложены принципы выбора сетевых параметров и управления шириной окна при нестационарных искажениях для нормальной процедуры обмена с групповым режимом отказа.

2. Открытые модели звена передачи данных

Предложена модель фрагмента сети [3–4], состоящего из двух последовательно соединенных звеньев передачи данных, в виде дискретной марковской системы массового обслуживания с конечным накопителем, отличающаяся от известных ранее совместным учетом трех факторов, которые определяют операционные характеристики линейного протокола управления межузловым соединением: искажений в прямом канале, искажений в обратном канале, блокировок буферной памяти транзитного узла. Анализ предложенной модели фрагмента сети показал, что его пропускная способность не превышает возможностей звена передачи данных с «худшими» параметрами. Найдена оценка оптимальной по критерию пропускной способности длины протокольного блока данных канального уровня, которая учитывает характеристики канала свя-

зи и параметры протокола ограничения на размеры буферной памяти, выделяемой для хранения пакетов данных в очередях к выходным каналам связи.

Для конвейерных процедур управления звеном передачи данных при дефиците буферной памяти показана целесообразность выбора ширины окна в размере объема всего буферного пула, выделенного выходному каналу связи, для режима селективного отказа – при любом качестве канала связи, а для режима группового отказа – при низком уровне искажений и высоких накладных расходах.

Установлено, что во всем диапазоне изменения достоверности передачи кадра предельные значения пропускной способности для всех режимов отказа достигаются уже при трех-пятикратном превышении объема буферного накопителя над размером окна. Для расчетов показателя пропускной способности селективного режима отказа управляющей процедуры можно использовать кусочно-линейную аппроксимацию, а для группового режима отказа – кусочное приближение.

3. Открытые модели многозвенного тракта передачи данных

Построены открытые модели многозвенного тракта передачи данных [5] с учетом фактора блокировок ограниченной буферной памяти транзитных узлов. Модели позволяют анализировать влияние емкости накопителей на показатели производительности сетевого протокола распределенной телекоммуникационной системы.

Обнаружена инвариантность показателя пропускной способности к порядку расположения транзитных узлов с буферными накопителями различной емкости вдоль статистически однородного тракта передачи данных и незначительная зависимость средней сквозной задержки пакета от этого порядка.

Установлена целесообразность равномерного распределения буферного пространства среди транзитных узлов вдоль многозвенного тракта, обеспечивающего наилучшую его производительность.

Показано, что при построении сетевых трактов передачи данных, состоящих из большого числа участков переприема, следует надежные каналы связи равномерно распределять между звеньями с высокими искажениями. Эти участки переприема выполняют роль дополнительных буферов между ненадежными звеньями и снижают отрицательный фактор блокировок буферной памяти.

Получены аналитическая оценка нижней границы пропускной способности и оценка сверху средней сквозной задержки многозвенного тракта передачи данных, соответствующие минимальному количеству буферов в транзитных узлах.

Предложена схема поиска распределения вероятностей состояний цепи Маркова, описывающей многозвенный тракт произвольной длины с единичными емкостями транзитных буферных накопителей, и его операционных характеристик.

4. Детерминированные конвейерные модели многозвенного виртуального соединения

Для различных условий функционирования неоднородной сети пакетной коммутации построены модели виртуального соединения с учетом влияния конвейерного эффекта [6–9], проявляющегося при пере-

даче мультипакетных сообщений по многозвенным виртуальным каналам, на время доставки пользовательских данных абонентам сети. Показано, что задержка мультипакетного сообщения в ненагруженном (пустом) виртуальном соединении в значительной мере определяется звеном соединительного пути с наибольшим временем передачи кадра.

Обнаружено свойство пространственно-временной симметрии детерминированного процесса информационного переноса, выражающееся в инвариантности показателя задержки одиночного однородного сообщения размера $N > 1$ в неоднородном виртуальном канале длины $D > 1$ и однородного сообщения размера $D > 1$ в однородном виртуальном канале длины $N > 1$ при равенстве элементарных задержек пакетов в виртуальных каналах. Задержка сообщения при этом инвариантна к расположению неоднородностей в пространстве (в тракте передачи данных) и порядку неоднородности во времени (в последовательности пакетов неоднородного сообщения).

Получено условие целесообразности фрагментации сообщения на пакеты при его передаче по многозвенному виртуальному соединению. Найдены аналитические зависимости оптимальной (по критерию минимума средней задержки абонентских сообщений) длины кадра от структуры сетевого трафика и параметров виртуальных каналов для идеальной однородной и неоднородной сети пакетной коммутации и ее оценки для сети с реальными свойствами каналов связи. Сформулировано правило выбора длины кадра, учитывающее ограничения рекомендаций ITU-T на размер информационной части пакета данных.

Для умеренной нагрузки на сеть развит метод выбора длины кадра и размера окна неоднородной сети передачи данных, отличающийся совместным учетом критериев системы (пропускной способности межузловых соединений) и пользователя (времени доставки пользовательских данных по виртуальным соединениям).

Из анализа задержки сообщения в нагруженном виртуальном соединении установлено, что при любом характере сетевого трафика (с пакетами одинаковой или различной длины) наилучшей в смысле времени доставки пользовательских данных стратегией формирования очередей к выходным каналам связи вдоль многозвенного тракта передачи является инверсное по длинам путей до адресата и размерам пакетов упорядочение элементов очереди.

При передаче по однородному виртуальному соединению однородного потока пакетов (пакетов одной длины) с его прямым переупорядочением в транзитных узлах задержка сообщения не зависит от состава отдельных очередей к выходным каналам связи (количества пакетов, адресованных в транзитные узлы).

Обнаружено свойство инвариантности задержки сообщения к структуре очереди в начале пути (порядку расположения пакетов в очереди, адресованных в различные узлы): для неоднородного виртуального соединения – при не возрастающих с длиной пути интервалах между передачей однородных пакетов по отдельным звеньям тракта передачи данных, а для однородного виртуального соединения с неоднородным трафиком – при неубывающем по времени передачи (длине) расположении пакетов в очереди перед сообщением для прямой и инверсной очереди.

5. Стохастические конвейерные модели протокола транспортного уровня

Построены стохастические модели процесса информационного переноса мультимедийных сообщений в многозвенном виртуальном соединении [10–12], отличающиеся учетом конвейерного механизма в передающем тракте с искажениями на отдельных участках переприема. Предложенные модели позволяют анализировать влияние длительности тайм-аута на операцию сквозной квитанции на операционные характеристики транспортного протокола.

Обнаружено свойство пространственно-временной симметрии стохастического процесса информационного переноса однородного потока пакетов в статистически однородном тракте передачи, проявляющееся в инвариантности вероятностно-временных показателей доставки данных удаленному абоненту к взаимно симметричным значениям размера сообщения $N > 1$ и длины виртуального канала $D > 1$.

Основной вклад в предельные значения средней сквозной задержки в стохастически однородном тракте, соответствующие неограниченной длительности тайм-аута, вносит время передачи мультимедийного сообщения и получения ответной квитанции в детерминированном конвейере при времени передачи в отдельной фазе, равном средней задержке пакета. Вклад остальных составляющих в сквозную задержку пропорционален интенсивности пакетных искажений, которыми для высококачественных каналов связи можно пренебречь.

Показано, что при двух-трехкратном превышении размера сквозного тайм-аута над заданной минимальной длительностью и низком уровне искажений в каналах связи однородного виртуального соединения для практических расчетов в большинстве случаев можно использовать соотношение для предельной задержки в стохастическом конвейере. Построен итеративный алгоритм выбора длительности тайм-аута ожидания сквозной квитанции по критерию заданного уровня вероятности повторной передачи данных.

6. Дискретные модели метода случайного множественного доступа

В настоящее время наиболее массовые технологии построения локальных вычислительных сетей (ЛВС) основаны на методе случайного множественного доступа к разделяемой среде передачи данных. Известные подходы к анализу производительности метода соперничества не различают в самой модели конфликта количества его участников. Проведенное исследование позволило изучить показатели производительности метода случайного множественного доступа в критических ситуациях, когда все станции ЛВС участвуют в конфликте. Предложена дискретная модель процесса соперничества, позволяющая определить зависимость пропускной способности и вероятностно-временных характеристик от количества информационных источников при заданной производительности моноканала [13].

Предполагается, что в ЛВС имеется K станций – источников данных, а время захвата среды передачи (моноканала) определяется слотовым интервалом (окном конфликта) длительности t . Считается, что все источники независимы, равноправны, всегда имеют блоки дан-

ных для отправки, одновременно входят в конфликт и в рамках процедуры соперничества используют механизм двоичной экспоненциальной отсрочки. Отсрочка выполняется в длительностях, кратных размеру слота t , а соперничество станций ЛВС продолжается до захвата моноканала одним из источников и выполнения успешной передачи. При выполнении повторных передач станции, обнаружившие конфликт других источников, продолжают соперничество в следующих повторных передачах согласно процедуре отсрочки. Количество повторных передач полагается неограниченным.

Для схемы группового участия всех информационных источников в конфликте найдены зависимости условной вероятности конфликта и условного среднего времени между последовательными конфликтами как функции номера повторной передачи и числа соперничающих источников. Из распределения параметров процесса соперничества следует, что с ростом количества станций ЛВС условная вероятность успеха в начальных повторных передачах снижается, однако при увеличении номера повторной передачи данная характеристика стабилизируется. Условное среднее время до повторной и успешной передач с ростом количества соперничающих вычислителей уменьшается, что легко объясняется тяготением вероятности условного успеха или неуспеха попытки захвата среды передачи данных в каждом цикле разрешения конфликта к началу периода отсрочки.

Анализ значений функции средней длительности соперничества выявил, что этот показатель монотонно увеличивается с ростом числа конкурирующих соперников. При этом составляющие данной характеристики – среднее время совокупного ожидания до начала повторных передач – практически инвариантны к количеству участников конфликта. К тому же среднее время, затрачиваемое на разрешение конфликтов и пропорциональное среднему числу столкновений на одну успешную передачу, монотонно растет с увеличением числа источников. Установлено, что среднее время совокупного ожидания до начала повторных передач практически не зависит от количества конфликтующих станций ЛВС, а верхняя граница среднего времени разрешения конфликтов при этом растет.

В реальной процедуре удвоение окна отсрочки прекращается после десятой попытки. Численные исследования алгоритма управления доступом к разделяемой среде с учетом этой поправки показали, что среднее время передачи кадра изменяется незначительно. Нормированная пропускная способность случайного метода доступа монотонно падает с ростом числа станций ЛВС и при равноправных источниках равномерно делится между ними, что приводит к пропорциональному сокращению индивидуального быстродействия для каждого источника информации. С ростом количества конфликтующих станций сети пропускная способность моноканала монотонно падает, средняя индивидуальная скорость передачи данных для каждого источника составляет от показателя пропускной способности долю, обратно пропорциональную числу соперников.

Анализ совокупности неоднородных по моменту вхождения в конфликт групп источников показывает, что с ростом номера группы (момента вступления в соперничество) условная вероятность захвата среды пере-

дачи данных станциями этой группы на N -й передаче существенно превышает данный показатель для источников с меньшим номером группы. Это преимущество тем больше, чем выше неоднородность групп (время между моментами вступления в конфликт) в силу значительного различия размеров периода отсрочки повторной передачи для станций различных групп. Источники, вступившие в конфликт самыми последними, имеют больше шансов захватить моноканал. Вероятность разрешения конфликта в пользу некоторой группы растет с увеличением ее объема (числа станций данной группы) и падает с ростом момента вхождения в процесс соперничества. Исследование значений среднего и среднеквадратического отклонения времени соперничества совокупности неоднородных групп источников установило, что наибольшее влияние на данные операционные показатели оказывает удаленность моментов вступления в конфликт различных групп от начала процесса соперничества. При вступлении группы в процесс соперничества после пятой повторной передачи вероятностно-временные характеристики системы практически не зависят от наличия этой и следующих за ней групп.

Рассмотрено влияние фактора необнуления счетчика повторных передач после успешной передачи одним из конфликтующих источников. Моделирование механизма независимого изменения счетчиков повторных передач для процесса соперничества произвольного числа станций свидетельствует о том, что любое отклонение счетчиков повторных передач от одинаковых значений приводит к снижению среднего времени соперничества.

7. Моделирование многоуровневой памяти вычислительных систем

Важнейшим элементом вычислительных систем является подсистема многоуровневой памяти, эффективность доступа к которой определяется степенью локализации приложений в кэш-памяти и адаптивными свойствами алгоритмов замещения кэш-строк. Степень локализации пропрограммного кода в значительной мере при прочих равных условиях определя-

ется коэффициентом ассоциативности кэш-памяти. На основе моделей многоуровневой памяти показано, что в кэше ассоциативности A с идеальным вытеснением при неравномерном распределении приложений $A - 1$ строк каждой группы содержат самые востребованные строки памяти. В то же время обнаружена инвариантность вероятности попадания в кэш к его ассоциативности при равномерном распределении приложений по пространству адресов.

Установлено, что условное среднее число сравнений признака кэша со старшей частью адреса объекта инвариантно к распределению приложения в памяти при последовательной реализации поиска кэш-строки в пределах ассоциативной группы. Обнаружено, что среднее время доступа к адресуемым объектам имеет линейную зависимость от степени локализации приложения в кэше. Показано, что с ростом глубины неблокируемости кэш-памяти пропускная способность многоуровневой памяти возрастает во всем диапазоне изменения вероятности попадания в кэш, а среднее время доступа к адресуемым объектам падает для широкой области высоких значений частоты попадания в кэш.

Для идеальной стратегии вытеснения самых неиспользуемых блоков в группе множественного ассоциативного кэша (имеющих наименьшие вероятности востребованности вычислителем) исследованы вероятностно-временные характеристики адаптивного движения процесса замещения блоков кэша к стационарному состоянию. Показано, что время движения распределено по геометрическому закону с меняющимся по мере достижения координат стационарного состояния параметром распределения.

Установлено, что минимальное среднее время движения к стационарному состоянию достигается для однородных вероятностей востребованности вычислителем блоков, отображаемых на заданную группу кэша. Исследовано влияние случайных отклонений от идеального процесса замещения блоков кэша на операционные характеристики многоуровневой памяти и динамические свойства кэш-памяти.

ЛИТЕРАТУРА

1. Сущенко С.П. Аналитическое оценивание оптимальных значений параметров линейного протокола сети ЭВМ с коммутацией пакетов // Автоматика и вычислительная техника. 1982. № 5. С. 66–71.
2. Сущенко С.П. Аналитические модели асинхронных процедур управления звеном передачи данных // Автоматика и вычислительная техника. 1988. № 2. С. 32–40.
3. Сущенко С.П. О влиянии блокировок буферной памяти на операционные характеристики звена передачи данных // Автоматика и вычислительная техника. 1985. № 6. С. 27–34.
4. Сущенко С.П. О влиянии блокировок буферной памяти на быстроедействие синхронных процедур управления звеном передачи данных // Автоматика и телемеханика. 1999. № 10. С. 115–125.
5. Сущенко С.П. О влиянии блокировок буферной памяти на производительность многозвенного тракта передачи данных // Автоматика и телемеханика. 1999. № 7. С. 66–79.
6. Сущенко С.П. Метод выбора рациональной длины пакета сети пакетной коммутации // Автоматика и вычислительная техника. 1984. № 3. С. 24–28.
7. Сущенко С.П. Параметрическая оптимизация сети пакетной коммутации // Автоматика и вычислительная техника. 1985. № 2. С. 43–49.
8. Сущенко С.П. Параметрическая оптимизация сети передачи данных с коммутацией пакетов // Проблемы и перспективы передачи и телеобработки данных. М.: Радио и связь, 1983. С. 107–108.
9. Сущенко С.П. Анализ сквозной задержки сообщения в многозвенном виртуальном канале // Автоматика и вычислительная техника. 1989. № 3. С. 52–64.
10. Сущенко С.П. Влияние длительности сквозного тайм-аута на задержку данных в виртуальном канале // Автоматика и вычислительная техника. 1991. № 6. С. 36–40.
11. Сущенко С.П. Анализ влияния длительности сквозного тайм-аута на операционные характеристики виртуального канала // Автоматика и вычислительная техника. 1995. № 4. С. 53–66.
12. Сущенко С.П. Аналитическая вычислимость сумм произведения показательных и степенных функций // Математическое моделирование и теория вероятностей: Сб. научных трудов ТГУ. Томск: Изд-во ТГУ, 1998. С. 253–256.
13. Кустов Н.Т., Сущенко С.П. Об устойчивости метода случайного множественного доступа // Математическое моделирование и теория вероятностей: Сб. научных трудов ТГУ. Томск: Изд-во ТГУ, 1998. С. 221–226.

Статья поступила в научную редакцию «Кибернетика и информатика» 15 декабря 1999 г.