

УДК 004.715

Е.В. Короткий, А.Н. Лысенко, д-р техн. наук

Поведенческая модель интегрального маршрутизатора

Рассмотрены функциональные блоки, архитектура и особенности конструкции маршрутизаторов для сетей на кристалле. Создана поведенческая модель интегрального маршрутизатора. Проведено исследование ее характеристик.

The survey of functional blocks, architecture and design features of network on chip routers is presented. The behavioral model of integrated router and the research on it is described.

Ключевые слова: *сеть на кристалле, интегральный маршрутизатор, System Verilog, поведенческая модель*

Введение

Одной из характерных тенденций, присущих современным системам на кристалле (СнК), является постоянное увеличение числа IP ядер при одновременном повышении требований к количеству и емкости межмодульных соединений [1]. Вследствие такого усложнения структуры традиционные способы организации подсистемы связи типа "общая шина" теряют эффективность [2,3]. Для решения проблемы обмена данными внутри сложных интегральных микросхем (ИМС) в работе [4] предлагается использовать концепцию сетей на кристалле (СтнК). Такой подход обладает преимуществами масштабируемости (с увеличением размера СтнК растет ее пропускная способность) и параллелизма (данные в различных областях сети передаются одновременно).

Общее представление о СтнК можно получить из ряда обзорных работ [1,5-6]. При этом значительную часть публикаций составляют теоретические исследования, в которых СтнК описывается при помощи аналитических либо высокоуровневых имитационных моделей. Исходный код существующих практических разработок в большинстве случаев не распространяется их авторами, а решения, имеющиеся в открытом доступе [7,8], как показано далее, обладают рядом недостатков. С другой стороны, теоретические исследования не должны быть оторванными от практики, поскольку именно практика позволяет выявить ограничения и область применения используемых методов и моделей.

Изложенное выше предопределяет актуальность выполнения практических разработок в

области СтнК и диктует необходимость создания аппаратной реализации ее компонентов. В данной статье рассмотрен этап описания поведенческой модели одного из узлов СтнК - интегрального маршрутизатора и проводится исследование ее характеристик.

1. Особенности конструкции компонентов интегральных сетей

СтнК имеют ряд отличий от своих макроаналогов из-за ограничений по доступным аппаратным ресурсам и потребляемой мощности [1]. Если компонент компьютерной сети может состоять из нескольких микросхем, то в случае СтнК все маршрутизаторы размещаются внутри одной ИМС. Данное обстоятельство также затрудняет использование больших объемов буферной памяти и сложных алгоритмов управления.

Вследствие применения коротких очередей ввода-вывода буферизация целого пакета становится нецелесообразной. Вместо этого используется так называемое "wormhole" управление потоком, когда пакет разбивается на атомарные единицы - флиты (flit – flow control unit), передаваемые непрерывно друг за другом [9]. Флиты продвигаются по мере возможности, не дожидаясь прихода последователей, что обеспечивает низкие требования к объему буферного пространства. Однако при таком подходе резко возрастает вероятность блокирования головы колонны (БГК) и, как следствие, возможно снижение пропускной способности маршрутизатора более, чем на 30% [10].

Использование механизма виртуальных каналов (ВК) является эффективным методом подавления БГК [11]. Его суть заключается в том, что физической линии связи ставится в соответствие несколько очередей, функционирующих параллельно. Для каждого порта определяется механизм доступа к ВК и способ арбитража запросов на прием-передачу. Блокирование одного из ВК не влияет на передачу пакетов через остальные каналы, позволяя избежать увеличения транспортной задержки. В то же время, необходимо принимать во внимание, что пропускная способность физического соединения разделяется между всеми ВК порта.

Поскольку большинство встраиваемых систем предназначено для выполнения одной за-

дачи, становится возможным уже на этапе разработки их структур определить потоки данных между модулями. Другими словами, пары приемо-передатчиков известны заранее, что позволяет упростить процедуру маршрутизации. Сложные алгоритмы определения маршрутов (такие, как OSPF и RIP) приводят к увеличению трафика в сети и требуют объемов памяти для хранения таблиц направлений [12]. Для устранения подобных недостатков в СтнК используется статическая маршрутизация от источника, когда сетевые интерфейсы хранят информацию о путях только к тем модулям, с которыми происходит обмен данными. В таком случае необходимость в таблицах направлений для маршрутизаторов исчезает, а информация о портах назначения передается в заголовке пакета [13].

Для обеспечения гарантированной доставки данных в классических сетях используется подтверждение приема, что приводит к увеличению нагрузки на сеть. В СтнК для этих целей применяется механизм “кредитов”, суть которого состоит в том, что количеству свободного места в буфере приемника соответствует некоторое число кредитов передатчика [13]. При освобождении ячейки во входной очереди количество кредитов передатчика инкрементируется, а в случае отправки флота - уменьшается на единицу. Продвижение пакетов возможно только при наличии у передатчика кредитов. Таким образом, гарантируется доставка без увеличения нагрузки на сеть.

2. Предлагаемая архитектура интегрального маршрутизатора и алгоритм его функционирования

На сегодняшний день в свободном доступе имеется только две реализации компонентов СтнК [7,8].

В работе [7] приводится описание маршрутизатора HERMES, созданного с использованием языка VHDL. Отсутствие поддержки ВК повышает вероятность блокирования пакетов и может привести, как уже упоминалось выше, к снижению пропускной способности узла более чем на 30% [10].

Работа [8] содержит сведения о конструкции интегрального маршрутизатора и сетевого интерфейса. Разработка оптимизирована под платформу Virtex-4 и не приспособлена для использования совместно с другими микросхемами FPGA. К недостаткам конструкции можно отнести отсутствие ВК и неэффективное использование маршрутной информации. Адрес и порт назначения передаются совместно с каждым флотом, что приводит к нерациональному применению около 10% соединений.

В данной публикации предлагается функциональная модель интегрального маршрутизатора, свободная от указанных выше недостатков. При ее создании использованы синтезируемые конструкции System Verilog, что должно максимально облегчить последующий переход на вентилярный уровень.

Структурная схема разработанного маршрутизатора содержит P входных портов, каждый из которых включает V ВК (рис.1). Величины P и V являются параметрами модели. ВК на входе подключаются к физическим соединениям на выходе при помощи коммутатора размерностью $P \times P$. В целях экономии памяти используется только входная буферизация, реализованная по принципу очереди. Гарантированная доставка данных обеспечивается применением механизма кредитов. Разрядность соединений является параметром модели и может быть произвольной.

Алгоритм функционирования маршрутизатора определяется модулем управления и состоит из следующих шагов, выполняемых последовательно:

а) Помещение флота в ВК. Каждый флот пакета содержит поле VC_ID с номером выделенного ему ВК. В момент прихода во входной порт флот помещается в очередь соответствующего ВК на основании анализа VC_ID . Разрядность VC_ID является параметром модели.

б) Определение порта назначения. Первые флоты содержат маршрутную информацию в виде последовательности портов назначения для каждого из промежуточных маршрутизаторов на пути следования пакета.

На этом этапе из головного флота извлекаются сведения о том, в какой выходной порт (физическое соединение) текущего маршрутизатора перенаправлять приходящие флоты, а ВК переходит в состояние “Установление соединения”. Данная операция выполняется один раз для каждого пакета. Более подробно используемый механизм маршрутизации (“от источника для сетей произвольного размера”) рассматривается в работе [13].

в) Выделение ВК. Пакету ставится в соответствие один из свободных ВК входного порта следующего маршрутизатора (выбранного на предыдущем шаге). В случае успеха полученное значение записывается в поле VC_ID всех флотов текущего пакета. Ситуация, когда свободные ВК отсутствуют, приводит к блокированию сообщения и запрос повторяется на следующем такте. Поскольку на свободные ВК следующего маршрутизатора могут претендовать каналы различных входных портов, при выполнении операции используется арбитраж. Различные методы арбитража

свободных ВК рассмотрены в работе [14]. В предлагаемой модели используется простейший из них [14], а его улучшение является предметом последующих исследований. Данный шаг выполняется один раз для всего пакета и в случае успешного завершения связь между ВК через физическое соединение считается установленной, а ВК текущего маршрутизатора переходит в состояние "Продвижение пакета".

в) Запрос на коммутацию входного ВК с одним из выходов маршрутизатора выполняется для всех флитов сообщения, когда соответствующий канал находится в состоянии "Продвижение пакета". Операция осуществляется для каждого ВК, который содержит данные для передачи, при условии наличия свободного места во входной очереди следующего маршрутизатора (кредитов передатчика). Размерность коммутатора равняется $R \times R$ и не зависит от числа V ВК. С одной стороны такой подход несколько повышает вероятность блокирования пакетов, а с другой - позволяет синтезировать решения, функционирующие на более высокой частоте [5]. Поскольку на каждое физическое соединение могут претендовать каналы различных портов, для выполнения операции используется арбитр [14]. Устройство управления конфигурирует коммутатор для установления связей, выигравших процедуру арбитража. При этом посылается кредит предыдущему маршрутизатору, информируя его о том, что во входной очереди освободилась единица места. В то же время число кредитов соединения со следующим маршрутизатором декрементируется. Если подтверждение пришло для последнего флита, ВК переходит в состояние "Свободен". В случае отрицательного ответа арбитра запрос повторяется на следующем такте.

д) Извлечение флита из очереди и передача через установленное соединение следующему маршрутизатору.

Описанные этапы могут выполняться на протяжении как одного, так и нескольких периодов тактовой частоты, что обуславливает нижнюю границу задержки маршрутизатора. Для разработанной модели минимальная задержка составляет два такта. В случае возникновения временных ограничений во время операции синтеза (следующий этап исследований), число тактов может быть увеличено, а процедура - конвейеризирована.

3. Моделирование и результаты исследований

На базе разработанного маршрутизатора создана СтнК с топологией в виде двумерной решетки размерностью 4×4 . С каждым узлом сети связано IP ядро генератора-приемника трафика,

имитирующее функционал вычислительного модуля и сетевого интерфейса. Тестовые программы описаны при помощи System Verilog.

Для определения промежуточных узлов на пути продвижения пакетов выбран метод поординатной XY маршрутизации, когда передача осуществляется вначале вдоль горизонтального направления, а затем вдоль вертикального [13]. Принятое решение обусловлено простотой генерации маршрутов при использовании данного метода. В случае необходимости можно использовать любой другой механизм маршрутизации.

Генераторы трафика создают пакеты, состоящие из 5-ти флитов и с интенсивностью λ вводят их в сеть. Интервалы времени между инъекцией распределены по экспоненциальному закону и в среднем равняются $1/\lambda$. Выбрано равномерно-случайное пространственное распределение сообщений, когда для каждого источника вероятность выбора того или иного приемника одинакова. При этом для каждого следующего пакета генерируется новый пункт назначения. На протяжении одного прогона модели все узлы генерируют по 1100 пакетов, а сбор результатов начинается после приема первых ста сообщений. Это осуществляется для обеспечения перехода СтнК в установившийся режим.

Моделирование проводилось в среде Modelsim 6.5 на персональном компьютере (Core2Duo 1.7ГГц, 1Гб ОЗУ) под управлением операционной системы Windows XP. Для автоматизации проведения исследования создан bash-скрипт, функционирующий в среде Cygwin 1.7.7 и осуществляющий серию прогонов модели для разных значений контролируемого параметра. Максимальное время одного прогона - 45 минут.

На рис.2 приведены результаты исследования по установлению связи между задержкой передачи СтнК и количеством ВК маршрутизаторов. В качестве задержки передачи принимается отрезок времени между инъекцией в сеть первого из флитов пакета и приемом его последней части в пункте назначения.

Как видно из рис.2, при низкой интенсивности введения пакетов транспортная задержка слабо зависит от числа ВК. Однако после увеличения приложенной нагрузки свыше 0.45 становится заметно, что наименьшую задержку и наибольшее значение порога насыщения (0.53) обеспечивает конфигурация с двумя ВК. Это обусловлено тем, что флиты пакетов передаются через физическое соединение поочередно. Следовательно, чем больше число ВК, тем выше период передачи флитов сообщения и задержка физического канала начинает доминировать над задержкой маршрутизатора.

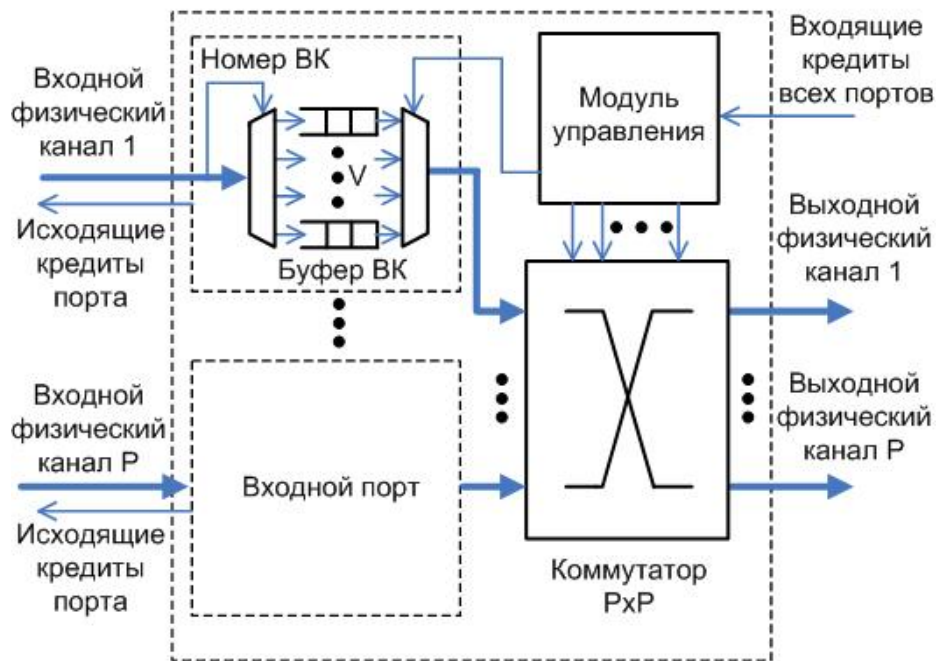


Рис. 1. Обобщенная структурная схема интегрального маршрутизатора: ВК – виртуальный канал, P – количество портов маршрутизатора, V – количество ВК для каждого порта

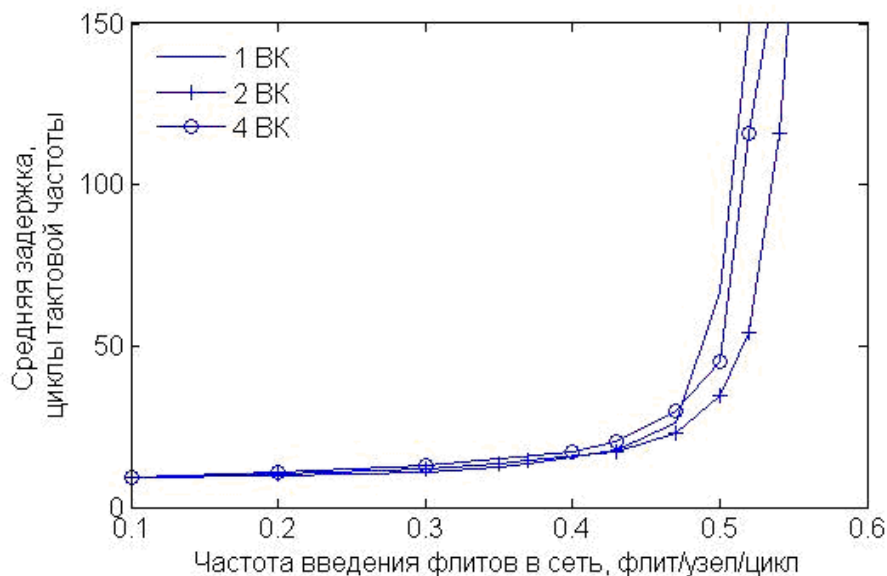


Рис. 2. Зависимость задержки передачи от приложенной нагрузки и количества ВК на входах маршрутизатора

Выводы

Анализ литературных источников позволил выявить основные отличия между СтНК и их макроаналогами. Для компонентов СтНК характерно использование коротких очередей, wormhole управления и простых алгоритмов маршрутизации. Механизм ВК применяется для снижения вероятности блокирования сообщений. Обозначенные меры обусловлены необходимостью уменьшения аппаратных ресурсов и мощ-

ности потребления при сохранении высокого порога насыщения СтНК. Приведенная авторами информация является полезной для разработчиков аппаратных реализаций компонентов СтНК.

Интегральные маршрутизаторы, имеющиеся в свободном доступе, имеют ряд недостатков, основными из которых являются отсутствие поддержки ВК и неэффективная передача маршрутной информации. В работе авторами предложена поведенческая модель маршрутизатора, свободного от указанных недостатков.

Использование данного решения позволит проводить эксперименты на более низком уровне абстракции по сравнению с аналитическими моделями и способствует повышению достоверности результатов исследований.

Анализ результатов моделирования показал, что увеличение числа ВК для соединений с высокой интенсивностью потока заявок приводит к росту порога насыщения сети и снижению транспортной задержки. С другой стороны, необоснованное повышение количества ВК способствует увеличению времени доставки. Ввиду обозначенных особенностей проведение разработок в направлении неравномерного распределения числа ВК для разных портов интегрального маршрутизатора представляется актуальным направлением последующих исследований.

Кроме того, поскольку представленная в работе процедура моделирования проводилась для равномерно-случайного пространственного распределения трафика, не характерного для реальных встраиваемых систем, предметом дальнейших исследований может быть проведение тестирования предложенной модели для конкретных практических приложений.

Литература

1. *Bjerregaard T., Mahadevan S.* A survey of research and practices of network-on-chip // *ACM Computing Surveys*. – 2006. – Vol.38, №1. – P.1-51.
2. *Angiolini F., Meloni P., Benini L.* A layout-aware analysis of networks-on-chip and traditional interconnects for mpsoCs // *IEEE Trans. on CAD of Integr. Circ. and Syst.* – 2007. – Vol.26, №3. – P. 421–434.
3. *Lee H.G., Ogras U.Y., Marculescu R.* On-chip communication architecture exploration: A quantitative evaluation of point-to-point, bus and network-on-chip approaches // *ACM Transactions on Design Automation of Electronic Systems*. – 2007. – Vol.12, №3. – P. 1-20.
4. *Dally W., Towles B.* Route packets, not wires: on-chip interconnection networks // *Proceedings of the 38th annual Design Automation Conference (June 2001)*. – Las Vegas, USA. – P.684-689.
5. *Atienza D., Angiolini F., Benini L.* Network-On-Chip Design and Synthesis Outlook // *Integration The VLSI Journal*. – 2008. – Vol.41, №3. – P.340-359.
6. *Marculescu R., Bogdan P.* The Chip Is the Network: Toward a Science of Network-on-Chip Design // *Foundations and Trends in Electronic Design Automation*. – 2009. – Vol.2, №4. – P.371-461.
7. *Moraes F., Calazans N.* HERMES: an infrastructure for low area overhead packet-switching networks on chip // *Integration, the VLSI Journal*. – 2004. – Vol.38, №1. – P.69-93.
8. *Ehliar A., Liu D.* An FPGA Based Open Source Network-On-Chip Architecture // in *Proc. of Fields Programmable Logic and Applications International Conf. (27-29 Aug. 2007)*. – Amsterdam, Netherlands. – P.800-803.
9. *Dally W.J.* Performance analysis of k-ary n-cube interconnection networks // *IEEE Transactions on Computers*. – 1990. – Vol.39, №6. – P.775-785.
10. *Karol M.J., Hluchyj M.G., Morgan S.P.* Input Versus Output Queueing on a Space-Division Packet Switch // *IEEE Transactions on Communications*. – 1987. – Vol.35, №12. – P. 1347-1356.
11. *Dally W.J.* Virtual-channel flow control // *IEEE Transactions on Parallel and Distributed Systems*. – 1992. – Vol.3, №2. – P.194–205.
12. *Олифер В.Г.* Компьютерные сети. Принципы, технологии, протоколы / Олифер В.Г., Олифер Н.А. – СПб.: Питер, 2007. – 958 с.
13. *Dally W.J.* Principles and Practices of Interconnection Networks / William James Dally, Brian Patrick Towles. – San Francisco: Morgan Kaufmann Publishers, 2004. – 550p.
14. *Dally W.J., Peh L.S.* A Delay Model and Speculative Architecture for Pipelined Routers // in *Proc. of 7-th Int. Symp. of High-Performance Comp. Arch. (20-24 Jan. 2001)*. – Nuevo Leone, Mexico. – P.255-266.