Т Е Х Н И Ч Е С К И У Н И В Е Р С И Т Е Т – В А Р Н А

инж. Милен Георгиев Ангелов

АРХИТЕКТУРА НА МАРШРУТИЗАТОР ЗА MPP И NUMA КОМПЮТРИ С DLH МРЕЖОВА ТОПОЛОГИЯ

А В Т О Р Е Ф Е Р А Т

на дисертация за получаване на образователна и научна степен "ДОКТОР"

Варна, 2018 г.

Дисертационният труд съдържа 173 страници и две приложения, включително 101 фигури и 5 таблици, оформени в 4 глави, общи изводи и списък на използваната литература от 130 заглавия, от които 1 на кирилица и 129 на латиница.

Защитата на дисертационния труд ще се състои на _____20___ г. от _____ ч. в _____ на открито заседание на жури сформирано със заповед на Ректора № _____ / _____ г.

Материалите по защитата (дисертацията, рецензиите и становищата) са на разположение на интересуващите се в Докторантския център, стая 318 НУК.

ХАРАКТЕРИСТИКА НА ДИСЕРТАЦИОННИЯ ТРУД

Актуалност на проблема

Всяка изчислителна система, без значение дали е суперкомпютър или персонален компютър, достига своята най-голяма производителност благодарение на използването на високоскоростни елементи и паралелно изпълнение на операциите. Възможността за паралелна работа на отделните изчислители и други модули във възлите на мултикомпютрите се явява един от главните фактори за увеличението на тяхната производителност.

Една от ясно изразените особености на съвременните изчислителни системи е наличието на комуникационни подсистеми, с чиято помощ паралелно работещите процесорни елементи осъществяват обмен на информация. Комуникационната подсистема е толкова важна за една изчислителна система, че много от характеристиките на нейната производителност се изразяват чрез термините за времената на междупроцесорен обмен. Това извежда на преден план комуникационните проблеми като особено актуални. Следователно, за повишаване производителността на една изчислителна система е необходимо да се реши задачата за синтез на ефективна комуникационна подсистема.

Съществува голям интерес относно мащабируемите паралелни системи с разпределена памет (Massively Parallel Processor – MPP), съставени от десетки и стотици хиляди изчислителни възли, свързани в мрежа. Те се явяват главно направление в развитието на компютърните архитектури с висока производителност. Удобството при тяхното мащабиране прави предаването на съобщения от предавател към приемник посредством комуникационни мрежи много по-приложимо, отколкото свързването на процесорите тип „обща шина", където тя е тясното място в комуникациите. Във връзка с това, че в MPP системите се осъществява пряка комуникация само между съседни възли, за предаването на съобщение от един към друг процесорен елемент, всеки от които е разположен в произволен възел от мрежата, е необходимо да се изпълни поредица от комутации на пакети.

Паралелните компютри, притежаващи до няколко десетки хиляди възела, имат своето приложение в различни области. Това е така, защото от една страна компютърните програми изискват високи скорости, които се постигат чрез паралелната работа при обработката на информацията, а от друга страна – повечето от тях използват само част от все по-големия брой процесорни ядра, изграждащи суперкомпютрите.

Проблем

Възлите в един паралелен MPP компютър са свързани във високоскоростна комуникационна мрежа. Връзките между съседните възли в паралелната система се реализират чрез маршрутизатори, които се използват за предаване на съобщенията по маршрутите им до тяхното местоназначение. Физическите канали между съседните маршрутизатори са жични и оптични връзки, реализирани в съответствие с избраната топология на свързаност. В тази връзка маршрутизаторите пряко участват в изграждането и мащабирането на системната

комуникационна мрежа и съществено влияят върху нейната производителност, като могат значително да намалят пропускателната способност на комуникационната подсистема на паралелните компютъри.

Цел и задачи на изследването

Цел: Създаване на архитектурна платформа на високоскоростен маршрутизатор за паралелен компютър, използващ DLH мрежова топология за реализация на своята комуникационна подсистема. Маршрутизаторът трябва да притежава висока пропускателна способност и ниска латентност.

Задачи:

1. Сравнителен анализ и избор на маршрутизация и техника за комутация за комуникационна мрежа с DLH топология. На тази основа да се разработи алгоритъм за маршрутизация, адаптиран към избраната мрежа.
2. Да се направи сравнителен анализ на каноничните архитектури за проектирането на маршрутизатори, използвани за изграждане на системните мрежи в паралелните компютри и да се направи избор на архитектурен модел за маршрутизатор, притежаващ необходимите качества. Да се структурират основните градивни елементи на данновите пътища в маршрутизатора - входни буфери, комутатор и изходни канали.
3. Да се анализират използваните алгоритми за разпределение и арбитраж на ресурси и да се направи избор на архитектура за основните функционални възли за управление на данновите пътища в маршрутизатора – разпределители, арбитри и контролери за физическите канали.
4. Да се проектира архитектурна платформа на маршрутизатор за високоскоростна статична DLH комуникационна мрежа, осигуряващ висока пропускателна способност, ниска латентност, липса на взаимна блокировка и висока степен на толерантност към откази.
5. Да се проектират алгоритмите за работата на основната част от функционалните възли на маршрутизатора - буфери, канали и арбитри.
6. Да се получат експериментални резултати за работата на разработените алгоритми, управляващи функционирането на основните блокове на маршрутизатора в тяхната взаимосвързаност, въз основа на които да се направят съответните изводи относно достигането на основната цел в дисертационния труд.

Обект и предмет на изследването

Обект на изследването е архитектурна платформа на маршрутизатор за паралелен компютър с DLH мрежова топология.

Предмет на изследването е работоспособността на функционалните блокове и алгоритмите в тяхната свързана съвкупност, както и латентността на пакетите, които се трансферират през маршрутизатор, изграден на базата на създадената архитектурна платформа.

Методи на изследване

За оценка на качествата и свойствата на основните видове маршрутизатори, които се използват за изграждане на системните мрежи в паралелните компютри, както и на техните функционални блокове, е използван методът на сравнителния анализ.

Като подход за изследване и оценка работата на функционалните блокове от проектираната архитектурна платформа на маршрутизатор, както и за оценка на латентността при трансфера на пакети през маршрутизатора, е използван методът на моделирането (симулация).

Място на изследване

Катедра КНТ при ФИТА на ТУ-Варна, България

Научна и практическа новост на изследването

Резултатите от проведените в съответствие с поставените цел и задачи на дисертационния труд изследвания се свеждат до следните основни приноси:

Предложена е концептуална платформа за проектиране на високоскоростен маршрутизатор за паралелни компютри, изградени на базата на DLH мрежова топология, осигуряващ ниска латентност, висока отказоустойчивост и липса на взаимна блокировка.

Разработен е алгоритъм за минимална адаптивна маршрутизация на пакети за комуникационна мрежа с DLH топология, подходящ за реализация чрез апаратни средства, чрез който само за един такт се определя посоката на предаване на един пакет.

Предложена е архитектура и е разработен алгоритъм на функциониране на входен буфер на маршрутизатора на базата на пул от FIFO опашки, директно свързани към комутационните канали и позволяващи в даден интервал от време от един и същ входен буфер към изходните канали да бъдат изпращани множество пакети за осигуряване на висока пропускателна способност. Подходът позволява достигане на пълно натоварване на изходните канали.

На основата на симулационни експерименти в среда Verilog при подходящо подбрани входни вектори, отразяващи критични за производителността на маршрутизатора ситуации, е доказана работоспособността на основните функционални възли на маршрутизатора и минималната латентност от два такта при трансфера на един пакет.

Апробация на изследването

Междинни резултати от изследванията са докладвани на четири международни конференции в България, публикувани в сборници с доклади „Компютърни науки и технологии” 2014 и 2015г., AUTOMATICS AND INFORMATICS’2014 и една статия в списание от 2017.

СЪДЪРЖАНИЕ НА ДИСЕРТАЦИОННИЯ ТРУД

ГЛАВА 1. ПРЕДАВАНЕ НА СЪОБЩЕНИЯ В ПАРАЛЕЛНИ МУЛТИПРОЦЕСОРНИ СИСТЕМИ

В първа глава са разгледани паралелните компютри и техните основни характеристики, свързани най-вече с организацията на паметта. По-задълбочен анализ е направен върху използваните топологии за свързване на възлите в една паралелна система и на комуникационните мрежи, чрез които се реализира обменът на данните. Извършеният анализ и сравнения позволиха да бъдат обобщени съответните изводи, както и да бъдат формулирани основната цел и задачите на дисертационното изследване.

Изводите, които са направени въз основа на анализа, представен в първа глава, са:

- Едно от най-важните изисквания при проектирането на паралелни компютри е постигането на балансирана системна производителност, формирана от паметта, процесорите, входно-изходните канали и комуникационната мрежа и ориентирането на тези компютри към широк кръг от научни и инженерни задачи. Хардуерът и софтуерът трябва да отговарят на нуждите на потребителите за много големи мащабируеми MPP системи за изчисления. От друга страна, тяхната изчислителна мощ трябва да бъде ефективно използвана от приложните програми, които да имат достъп до най-големия капацитет от компютърни ресурси или до по-малка част от тях в зависимост от конкретните изисквания и условия при разпределяне натоварването на системата.
- Всеки паралелен компютър използва комуникационна подсистема за свързване на процесорите, паметите и дисковете, които използват системни и входно/изходни магистрали, както и свързаните чрез някакви интерфейси процесори към локалните мрежи. За съжаление, комуникационната подсистема почти винаги става тясното място за повечето приложения при паралелните компютри. Недостатъчното бързодействие и ограничените комбинационни възможности могат съществено да снижат ефекта от паралелната работа на множеството процесори в системата. Изборът на топология и на алгоритъм за маршрутизация на пакетите са от голямо значение за пропускателната способност на комуникационната мрежа.
- Водещите компании – производители на суперкомпютри обръщат много голямо внимание при проектирането и реализацията на системните комуникационни мрежи и на модулите, които ги изграждат и осъществяват обмена на информация между отделните възли на системите.
- Маршрутизаторите във възлите на един MPP или NUMA компютър са модулите, които реализират предаването на данните до тяхното местоназначение. В един паралелен компютър с обмен на съобщения те формират предаването на пакетите по техните маршрути. Каналите между съседните маршрутизатори и топологията на свързаността може да варира от система към система. Ясно е, че в една мрежа те могат да бъдат пречка за мащабирането и съществено да намалят нейната производителност.

- Качествата на Double-Loop Hypercube (DLH) мрежовата топология комбинират предимствата на класическата хиперкубична топология - малък диаметър, висока свързаност при невисока структурна сложност (сравнително малък брой връзки, реализирани от всеки възел), симетричност, отказоустойчивост и проста маршрутизация от една страна и от друга страна - мащабируемост на мрежата и постоянна степен на свързаност на възлите. Диаметърът на мрежата зависи от размерността на хиперкубовете, които я изграждат и от техния брой. Мащабируемостта на една система, използваща DLH топология, може да се реализира лесно чрез увеличаване броя на хиперкубовете. Разпределението на задачите и на ресурсите се прави от операционната система. В общия случай една задача може да заеме възлите на част от един хиперкуб, цял хиперкуб или необходимия брой хиперкубове от цялата мрежа, които може да й предостави операционната система. Ясно е, че използването на съседни възли (подкубове) или съседни хиперкубове от една задача ще намали броя на хоповете, а следователно ще намали и латентността при обмена на съобщения между възлите, необходим при нейното решаване.
- Паралелните компютри с няколко десетки хиляди възела, притежаващи все по-голям брой процесорни ядра, имат широко приложение в различни области. В това отношение интерес представлява реализацията на ефективен обмен на данни в мултипроцесорни системи, чиято свързаност е изградена на базата на комуникационни мрежи с DLH топология и маршрутизаторите, лежащи в основата на този обмен.

ГЛАВА 2. **АНАЛИЗ НА АЛГОРИТМИ И СТРУКТУРИ, ИЗПОЛЗВАНИ В МАРШРУТИЗАТОРИТЕ ЗА ПАРАЛЕЛНИ МУЛТИПРОЦЕСОРНИ СИСТЕМИ**

Във втора глава е направен сравнителен анализ на структури, методи и алгоритми, използвани при проектирането на маршрутизатори за системни мрежи на паралелни компютри. На тази база е направен мотивираният избор на определящите принципи, структури и алгоритми, които са в основата на проекта.

2.1. Маршрутизация: Алгоритъмът за маршрутизация, използван в една мрежа, е от решаващо значение поради няколко причини: 1). Балансиране на натоварването в каналите на мрежата дори при модел на нееднороден трафик; 2). Поддържане на възможно най-кратък път за един пакет, намалявайки броя на хоповете и цялостната латентност на пакета, използващ този път; 3). Работа в присъствието на откази и неизправности в мрежата.

При избраната DLH топология на мрежата и направения сравнителен анализ на видовете маршрутизация, в настоящия проект се използва минимална адаптивна маршрутизация, защото: 1.) Осигурява избор на маршрут с минимален брой хопове с цел минимална латентност; 2.) Осигурява възможности за алтернативни минимални маршрути, в зависимост от моментното състояние на мрежата (натоварване, дефектирали възли и канали за връзка); 3). Не води до зацикляне на пакети в мрежата; 4). Регулярната структура на DLH топологията позволява реализация на алгоритмична маршрутизация във всеки възел.

2.2. Комутация: Техниките за комутация реализират преместването на информационните единици вътре във всеки маршрутизатор напред по пътя от източника до получателя. За сравнение на анализираните техники за комутация е използвана латентността на един пакет от $\boldsymbol{m}$ бита в отсъствието на трафик. Ясно е, че реалните латентности зависят от безброй подробности на реализацията и условията за обмен. За еднаквост на условията за сравнението се приема, че: 1). Размерът на един флит е равен на размера на един фит и на физическата ширина на канала от $\boldsymbol{W}$ бита; 2). Заглавната маршрутна част на пакета е един флит, дължината на съобщението е $\boldsymbol{m+W}$ бита и един маршрутизатор може да вземе решение за маршрутизация за $\boldsymbol{t_r}$ секунди; 3). Физическият канал между два маршрутизатора работи с честота $\boldsymbol{B}$ Hz и пропускателната способност на канала е $\boldsymbol{WB}$ bps при условие, че предаването на един фит се реализира за един такт. Тогава времето за разпространение в този канал е $\boldsymbol{t_w=1/B}$; 4). След като един маршрут е бил създаден веднъж през маршрутизатора, вътрешното забавяне, или времето за комутация на един флит е $\boldsymbol{t_s}$; 5) Терминалните възли в мрежата са на разстояние $\boldsymbol{h}$ линка един от друг. Тогава за латентността се получава:

- при комутация на канали: $t_{circuit}=t_{setup}+t_{data}$ (2.1)
  $t_{setup}=h[t_r+2(t_s+t_w)]$ (2.2)
  $t_{data}=(m/W)/B$ (2.3)
- при SF комутация на пакети: $t_{packet}=h[t_r+(t_s+t_w)(1+ m/W)]$ (2.4)
- при VCT метод за комутация: $t_{vct}=h(t_r+t_s+t_w)+max(t_s, t_w)(m/W)$ (2.5)
- при WH метод за комутация: $t_{wh}=h(t_r+t_s+t_w)+max(t_s, t_w)(m/W)$ (2.6)

2.3. Архитектура на маршрутизатор: Един маршрутизатор се състои от регистри, комутатори, функционални блокове и логика за управление, които съвместно осъществяват маршрутизацията и функциите за управление на потока от входните буфери напред по пътя към техните дестинации. Заглавните флитове преминават през степените на конвейерите, които извършват маршрутизация и разпределение на виртуалните канали, след което всички флитове преминават комутатора и изходните устройства. Архитектурата на един маршрутизатор в голяма степен се определя от техниката на комутация, която той реализира. Повечето от съвременните маршрутизатори за паралелни мултипроцесорни системи използват някаква форма или вариант на VCT, WH или буферирана WH техника на комутация. Архитектурата е свързана с топологията на комуникационната мрежа, буферите и техниката на комутация, имплементирана в един маршрутизатор. Основните цели са постигане на едно или няколко от следните качества: ниска латентност на пакетите, отказоустойчивост на мрежата, ниска консумация на енергия, ниска цена и възможност за реконфигуриране на някои от параметрите в маршрутизатора, което дава гъвкавост при неговото използване.

2.4. Даннови пътища в маршрутизаторите: Пътят за данните в един маршрутизатор манипулира съхранението и движението на данновите пакети и се състои от набор входни буфери, комутатор и изходно устройство. Останалите блокове изпълняват управлението в маршрутизатора и са отговорни за координиране движението на пакетите чрез средствата на техния даннов път. Обикновено блоковете за управление изпълняват изчислението на маршрутите, разпределението на виртуалните канали и разпределението на каналите в комутатора:

2.4.1.Входни буфери: Състоят се от памет, организирана по определен начин и нейното управление. Протоколът за управление на данновия поток разпределя пространството на паметта във входните буфери за съхраняване на флитовете, очакващи възможност за напускане на маршрутизаторите. Разделянето на входните буфери е тясно свързано с проектирането на комутатора. Те могат да бъдат централизирани, разпределени към физическите канали, или разпределени към виртуалните канали. Паметта за всеки един буферен дял може да бъде статично разпределена към всеки буфер на виртуален канал, като се използва механизма на кръговите буфери, или да бъде динамично разпределена с помощта на свързан списък;

2.4.2. Комутатор: Той е основен компонент в един маршрутизатор. Неговото предназначение е да преведе насочените пакети от флитове от входните до техните целеви изходни портове;

2.4.3. Изходни устройства: Данновият път в изходния блок на маршрутизатора се състои основно от FIFO буфер, чрез който се формира съответствие относно скоростта на предаване на изходния канал към скоростта на предаване през комутатора. За комутатор, който няма изходно ускорение не се изисква FIFO буфер. Когато флитовете напускат комутатора, те могат да бъдат поставени директно върху изходния канал.

2.5. Блокове за управление в маршрутизаторите: Управляващите блокове на маршрутизаторите до голяма степен се състоят от арбитри и разпределители. Те определят данновия път за всяка една от пристигналите даннови единици (пакети и флитове):

2.5.1. Арбитри: Определят избора на една от множество заявки за един ресурс. Те формират основния градивен елемент за разпределителите, които съпоставят множеството заявки с множеството ресурси;

2.6. Разпределители: Определят виртуалните канали за пакетите и разпределят циклите за комутация на флитовете или на пакетите. Докато арбитърът присвоява един ресурс към един източник от група източници на заявки, разпределителят реализира съвпадение между група от ресурси и група от източници на заявки, всеки от които може да поиска един или повече от ресурсите.

2.7. Управление на потока в маршрутизаторите: синхронизиран протокол за предаване и приемане на информационни единици. Единицата за управление на потока се отнася до тази част на съобщението, чийто трансфер трябва да бъде синхронизиран и се дефинира като най-малката единица информация, чието прехвърляне се изисква от подателя и се потвърждава от приемника. Сигнализацията за заявка / потвърждение се използва за гарантиране на успешен трансфер и за наличието на буфери в приемника.

2.9. Изводи: На база на извършените сравнения и анализи са направени следните изводи:

- Висока пропускателна способност на комуникационната система може да бъде постигната чрез подходяща архитектура на маршрутизаторите в съответствие с избраната топология на системната мрежа, използвания алгоритъм за маршрутизация и протокол за трансфер. Целта е да се удовлетворят изискванията за ниска латентност на пакетите, мащабируемост и отказоустойчивост на мрежата, гъвкавост при използване и разумна сложност на реализация.
- Важни компоненти в архитектурата на един маршрутизатор, които съществено влияят върху неговата производителност са степента на паралелизъм, реализацията на процесите на маршрутизация и арбитраж, техниката на комутация на входните към изходните канали и буферирането на пакетите.
- Архитектурата на един маршрутизатор в голяма степен се определя от техниката на комутация, която той реализира. Наложили са се архитектури с Wormhole и Virtual Cut-Through методи за комутация с използване на виртуални канали, които са сравними по отношение на бързодействието. Cut-Through техниката изисква по-големи буфери, но постига максимална скорост при последователното предаване флитовете на един пакет и елиминира възможност за взаимна блокировка на пакетите.
- В съответствие с поставената основна цел и на базата на направения анализ на градивните блокове в един маршрутизатор за паралелен компютър, е направен изборът за:

- Архитектура на маршрутизатора: Пряко свързване на входните буфери към комутатора, което намалява латентността на пакетите и позволява свързване на повече от една опашка от един буфер към изходите. Арбитраж при всеки от изходните канали, което води до решаване на конфликти само при изходите и максимално използване на пропускателната способност на маршрутизатора. Използване на Cut-Through техника за комутация, което намалява латентността на пакетите, тъй като се елиминират контролерите на виртуалните канали, свързани с изходите. Премахва се опасността от deadlock блокиране поради „прибиране" на целия приеман пакет в текущо използвания входен буфер и освобождаване на останалите ресурси, използвани от този пакет;

- Входен буфер: Пул от FIFO опашки, всяка от които може да съхрани всички флитове от един текущо приеман пакет. Това осигурява максимална скорост при приемане на флитовете и дава основа за премахването на опасността от deadlock блокиране;

- Комутатор: Тип Crossbar с независимо мултиплексиране на входните опашки към изходите;

- Изходен канал: Наличие на кръгов арбитър за справедлив избор на входна опашка, която да бъде свързана към изхода. Това осигурява висока пропускателна способност на канала и на маршрутизатора като цяло. За буфериране на изхода се използва само регистър, което води до намаляване на апаратната част, спестяване място на кристала и намалена консумация на енергия;

- Арбитраж при изходите: Използване на кръгов арбитър с два паралелно работещи PPT_Pre_Thermo блока. Тази архитектура постига справедлив арбитраж и висока скорост при конвейерните маршрутизатори с виртуални канали без конфликти при изходите (без предварително мултиплексиране на FIFO опашките към комутатора);

- Разпределение: Използване на двустепенен разпределител с iSLIP алгоритъм, изграден от арбитри при входните опашки и при изходните канали. Разпределението се извършва с една итерация във всеки времеви интервал.

- Изборът на архитектура и използваните принципи за изграждане структурите на основните блокове дават основата за създаване на архитектурна платформа на маршрутизатор, притежаващ висока пропускателна способност и ниска латентност.

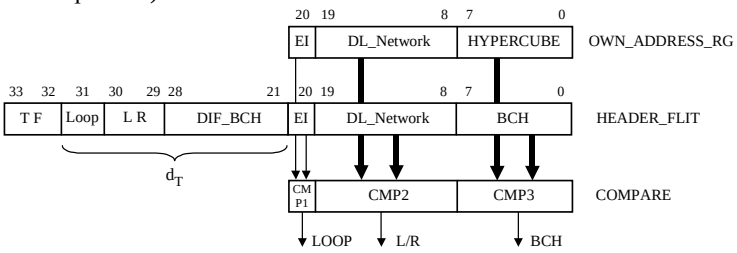
ГЛАВА 3. **МАРШРУТИЗАТОР ЗА ПАРАЛЕЛНИ МУЛТИПРОЦЕСОРНИ СИСТЕМИ С DLH МРЕЖОВА ТОПОЛОГИЯ**

В трета глава в най-общ вид са представени взаимно обосновани технически и архитектурни решения, използвани при проектирането на архитектурната платформа на маршрутизатор за паралелен компютър с DLH мрежова топология.

3.1. Принципи, изисквания и ограничения: В тази точка са формулирани основните принципи, изисквания и ограничения, използвани при създаването на архитектурната платформа.

3.3. Алгоритъм за маршрутизация: Разработеният алгоритъм избира между минималните маршрути от възел-източник до възел-получател, като използва информация за състоянието на мрежата при вземането на решение за маршрутизация на всяка стъпка. Функцията за маршрутизация генерира продуктивен изходен вектор, който определя чрез кои изходни канали на съответния възел ще може да се премести приетия пакет по-близо до неговия получател. Основни предимства: 1). Реализира избор измежду няколко пътя с минимална дължина, което се позволява от мрежовата топология; 2). Позволява бърза схемна реализация в комбинация с фазата на арбитраж в конвейера на един маршрутизатор; 3). Състоянието на мрежата се определя по текущия статус на опашките във входните канали на съседните маршрутизатори; 4). Не води до зацикляне на пакети в мрежата.

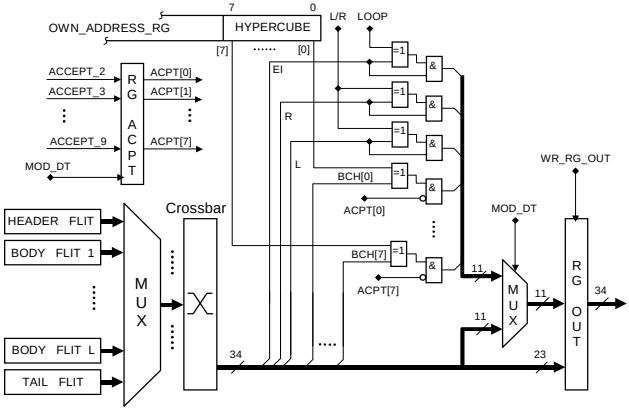
Алгоритъмът използва информацията от заглавния флит на преминаващия пакет и резултата от сравнението му с адреса на текущия възел, записан в специален регистър OWN_ADDRESS_RG в маршрутизатора. Сравнението се прави във всеки възел от пътя на един пакет от източника до получателя. Схемата за сравнението е показана на Фиг.3.1. Високо ниво на резултантен сигнал - LOOP, L/R или BCH показва равенство на съответните сравнявани полета от адреса (EI - бит, указващ външен/вътрешен кръг от DL мрежата, DL_Network – целеви адрес на хиперкуб в DL мрежите, BCH – целеви адрес на възел в избран хиперкуб от DLH мрежата).



Фиг.3.4. Схема за сравнение адреса на текущия възел с целевия адрес на пътуващ пакет

1). Ако DEST_ADR=OWN_ADR, то това е възелът - получател и от активната опашка на входния канал се подава заявка към изходния канал на собствения възел за приемане на пакета; 2). Ако DEST_ADR ≠ OWN_ADR, към арбитрите на изходните канали се изпращат толкова заявки, колкото са активните битове от

полето за несъвпадение d_T. След успешно разпределение за пакета и установяване на път през комутатора на текущия маршрутизатор, полето d_T се модифицира в зависимост от избрания изходен канал, преди заглавния флит да бъде изпратен към следващия по веригата маршрутизатор чрез схемата, показана на Фиг.3.5.



Фиг.3.5. Схема за модификация на полето за несъвпадение в заглавния флит

3.4. Основни функционални блокове на маршрутизатора

На Фиг. 3.6 е показана проектираната обща блокова схема на маршрутизатор за паралелен компютър, свързан в DLH(12,8) мрежа със следните основни блокове:

- INITIALIZER: Блок за инициализация. Свързан е към HOST процесора на възела чрез данновите линии на входния канал на собствения възел. Участва в работата на останалите функционални блокове;

- SWITCH: Crossbar комутатор, свързващ входните опашки с изходите;

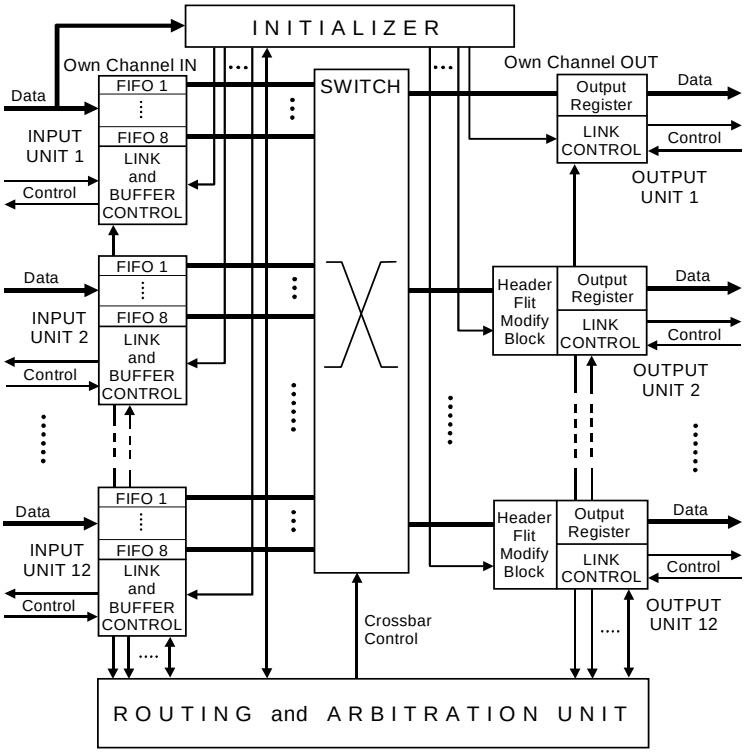
- INPUT UNIT: Блокът е съвкупност от взаимосвързани модули, реализиращи приемането на данни и управлението на един входен канал и включва: входен буфер с неговия пул от опашки, управляващите автомати на буфера и на всяка от опашките, логически схеми, входно/изходни буфери за сигнали и др.;

- OUTPUT UNIT: Съвкупност от взаимосвързани модули, които реализират предаването на данни и управлението за един изходен канал. Блокът включва: изходен регистър, логически блок за модифициране на заглавния флит (без OUTPUT UNIT за собствения канал), управляващ автомат на изходния канал и неговия арбитър, логически схеми, входно/изходни буфери за сигнали и др.;

- ROUTING and ARBITRATION UNIT: Блокът реализира разпределението на входовете към изходите след успешен арбитраж. Включва входни и изходни арбитри и логически схеми;

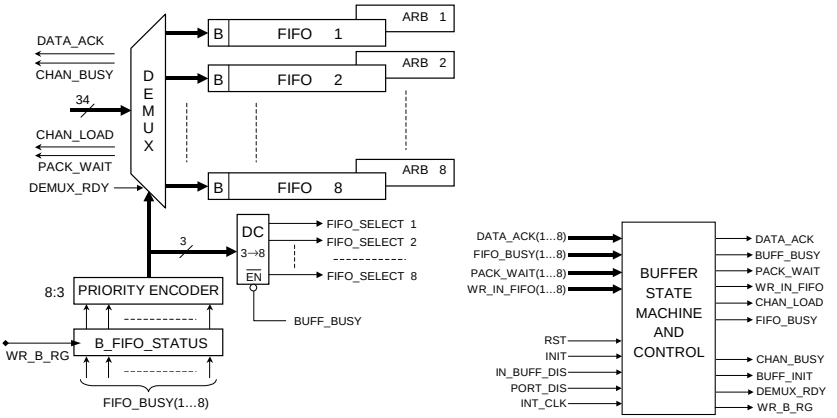
Връзките между отделните блокове са показани чрез два типа шини: даннови (означени с дебели линии) и управляващи означени с по-тънки линии). Стрелките на линиите показват посоките на трансфер на данните и въздействието на управляващите сигнали.

3.4.2. Входен буфер: На Фиг.3.9 е показана проектираната архитектура на входния буфер на маршрутизатора, използващ Cut-Through техника за комутация,

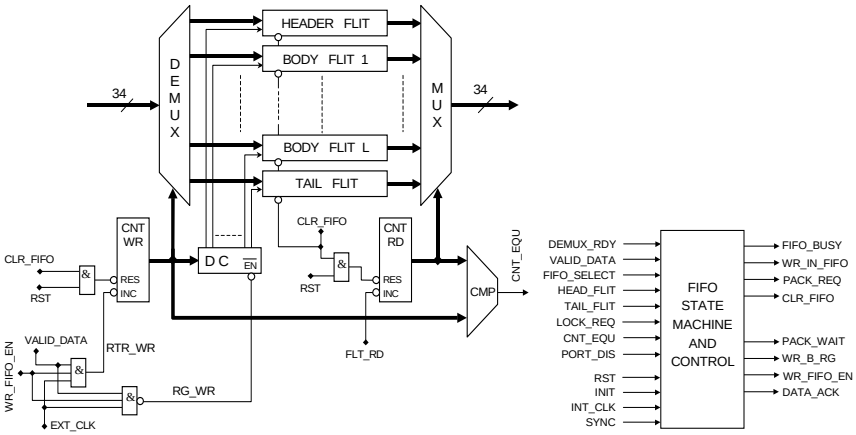


Фиг.3.6. Архитектурна блокова схема на маршрутизатор за паралелен компютър

осигуряваща даннов обмен на ниво пакет. Буферът се състои от: демултиплексор DEMUX, FIFO опашки, приоритетен енкодер PRIORITY ENCODER, регистър за състоянията на опашките B_FIFO_ STATUS, декодер за избор на FIFO опашка DC и блок за управление. Основата му е пулът от осем FIFO опашки, чиято структура е показана по-долу на Фиг.3.10. Всяка опашка има свой локален арбитър ARB[k], k=1..8, който при необходимост генерира заявки за наличие на заглавен флит към кръговите арбитри на изходните канали. Една опашка може да съхрани само един пакет с дължина, не по-голяма от броя на нейните запомнящи елементи. Състои се от: демултиплексор DEMUX, (L+2) броя запомнящи елементи за флитове, брояч за запис CNT_WR в комплект с декодер DC за селектиране на пореден запомнящ елемент, брояч за четене CNT_RD, компаратор CMP, мултиплексор MUX, логика и управление. Ширината на запомнящите елементи е 34 бита [D_{33}..D_0]. Ширината на един флит е 32 бита [D_{31}..D_0]. D_{33} и D_{32} кодират типа на флита: HEADER, BODY или TAIL. Всяка от опашките може да бъде в едно от две състояния: заето и свободно, показано на Фиг.3.9 чрез флага B (Busy). Състоянието на входния канал в зависимост от състоянията на опашките може да бъде: I). Зает (CHAN_BUSY=1, всички опашки са заети); II). Силно



Фиг.3.9. Архитектура на входен буфер на маршрутизатор за паралелен компютър

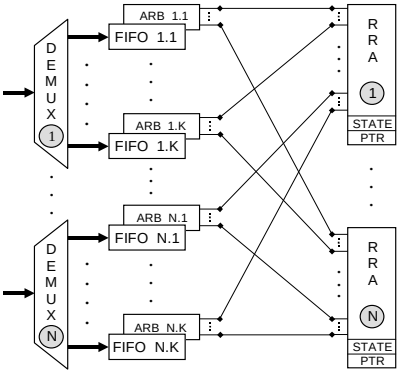


Фиг.3.10. Архитектура на FIFO опашка на входен буфер на маршрутизатор

натоварен (CHAN_BUSY=0, CHAN_LOAD =1, свободна е само една опашка); III). Слабо натоварен (CHAN_BUSY=0, CHAN_LOAD=0, свободни са две или повече опашки). Това състояние се предава към съседния маршрутизатор чрез сигналите CHAN_BUSY и CHAN_LOAD. Приоритетният енкодер винаги избира първата свободна опашка. Тя може да приеме всеки пристигащ пакет, без значение за кой изходен канал е предназначен. След началото на приемане на пакет избраната опашка преминава в състояние заето (B=1) и подава сигнал PACK_REQ (показан на Фиг.3.3) към локалния арбитър, който стартира фазата за маршрутизация и арбитраж.

3.4.3. Блок за маршрутизация и арбитраж

На Фиг.3.13 е показана в общ вид архитектурата на разпределителя на маршрутизатора, който е основата на проектирания блок за маршрутизация и



Фиг.3.13. Архитектура на двустепенен разпределител на маршрутизатор

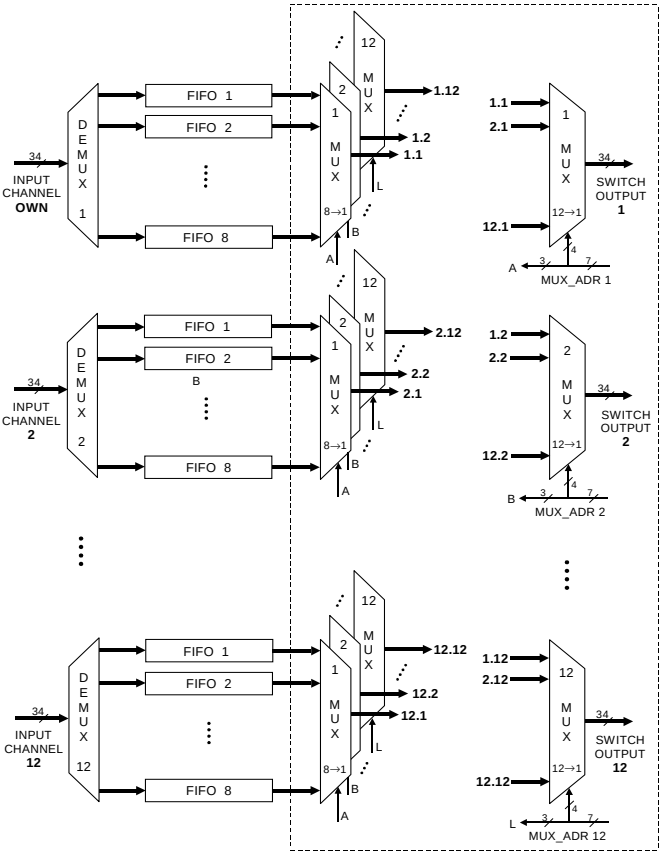
арбитраж, показан на Фиг.3.6. За яснота освен арбитрите, са показани демултиплексорите и опашките на входните канали. Разпределителят се състои от две групи арбитри: входни и изходни. Алгоритъмът за маршрутизацията е реализиран апаратно, като основната част от него е имплементирана в арбитрите и тяхното управление. Маршрутизаторът има *N=12* броя входни буфера, всеки от които съдържа *K=8* броя FIFO опашки и *N=12* броя изходни канала. Всеки работещ възел от компютърната система съдържа информация за дефектиралите възли и линии за връзка от мрежата, която се взема задължително в предвид при формиране маршрутите на пакетите и се зарежда при всяка инициализация в маршрутизатора.

3.4.4. Комутатор

Проектираният комутатор е показан на Фиг.3.18. Той осигурява директна свързаност и реализира мултиплексирането на всяка от входните FIFO опашки към всеки от изходните канали на маршрутизатора. Освен компонентите на комутатора, визуално групирани в правоъгълното поле, обозначено с пунктирана линия, на същата фигура са показани демултиплексорите и опашките на входните канали. Самият комутатор е тип Crossbar и е реализиран чрез каскадно свързани мултиплексори. Той има 96 входни и 12 изходни порта, всеки с ширина 34 бита, следователно всеки канал на мултиплексорите също има ширина 34 бита. Адресите за всяка една от групите мултиплексори, обслужваща един изходен порт, се генерират от RRA арбитър на разпределителя, предназначен за съответния изходен канал.

3.4.5. Изходен канал

Както е показано на Фиг.3.6, изходния канал на маршрутизатора се състои от паралелен регистър-памет Output Register, логически блок за модификация на заглавния флит Header Flit Modify Block и управление LINK CONTROL. Арбитърът на изходния канал е включен в блока на разпределителя, а неговото управление се извършва от блока LINK CONTROL. От архитектурна гледна точка Output Register и буферите на изходите не представляват интерес. На Фиг. 3.5 е показана проектираната схема за модификация на полето за несъвпадение, която пряко участва в маршрутизацията на пакетите. Когато заглавният флит



Фиг.3.18. Архитектура на комутатор с директна връзка на входните опашки към изходите

минава по вече установения маршрут, неговото поле d_T се модифицира с помощта на полето HYPERCUBE от OWN_ADDRESS _RG, съдържанието на регистъра RG_ACPT (в него се съхранява кой е избрания изходен канал към следващия по маршрута на пакета възел от хиперкуба), резултатите от сравнението LOOP и L/R, логики и мултиплексор. Резултатът от текущата модификация е нулиране на онзи бит от полето d_T (при определени условия), чрез който е избран изходния канал за пакета в текущия възел. Така на всеки хоп се намалява броят на заявките към пожеланите изходни канали в следващия възел от маршрута, което предотвратява и зациклянето на пакета в мрежата. Сигналите, които управляват модифицирането на флита са: MOD_DT (управление на мултиплексора) и WR_RG_OUT (записва модифицираното поле d_T в регистъра RG_OUT на изходния канал, където заглавния флит става достъпен за приемане от следващия по веригата маршрутизатор). Всички останали флитове от пакета преминават без модификация.

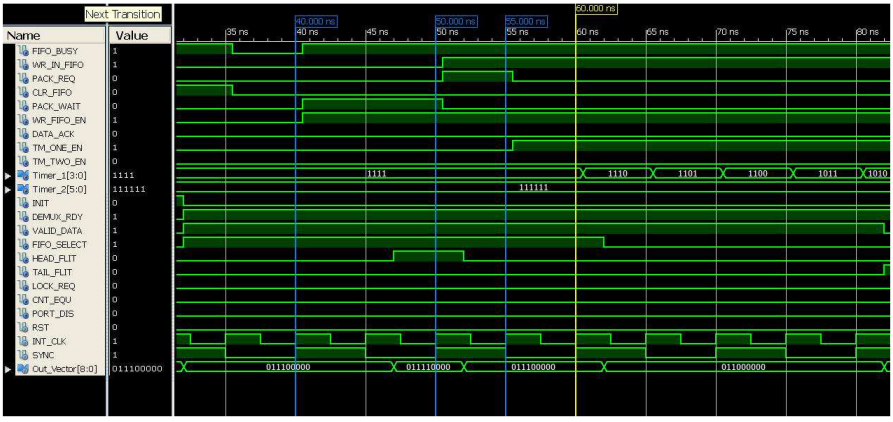
ГЛАВА 4. **ЕКСПЕРИМЕНТАЛНИ ДАННИ И РЕЗУЛТАТИ**

В тази глава са представени експериментални резултати от изследванията, направени за управлението работата на проектираните основни блокове, от които се състои архитектурната платформа на маршрутизатора. Изследванията са извършени чрез симулации с помоща на програмен продукт PlanAhead, v14.6, на фирмата Xilinx. Програмите, описващи работата на устройствата, са написани на Verilog. Симулациите изследват и визуализират чрез времедиаграми работата на отделните функционални възли и тяхната взаимосвързаност в една система. За извършване на изследванията периодът на вътрешния тактов сигнал INT_CLK в маршрутизатора е избран да бъде 5.00ns, а закъснението при превключване на състоянията DLY да бъде 0.50ns. Тези две стойности се използват основно за визуализацията на сигналите върху времедиаграмите и за постигане на по-голяма яснота при обясненията по-долу. Самите резултати се оценяват в брой тактове Показаните времедиаграми са за малка част от изследванията, показваща съществената част от резултатите. Също така, за яснота при обясненията се използва наименованието “текущ маршрутизатор”, който работи в система, свързан с физически канали към неговите „съседни маршрутизатори”.

4.1. Експериментални резултати за блока за инициализация: В тази глава са изследвани са различни комбинации от възможни входни последователности от сигнали, които доказват правилното поведение на блока в съответствие със заложените изисквания към неговата работа.

4.2. Експериментални резултати за блока за инициализация: В тази глава е доказана правилността на работата на проектираните входни буфери на маршрутизатора, управляващи своите пулове от FIFO опашки за пристигащи пакети и тяхното управление. Изследвани са възможните състояния и въздействието на входните сигнали върху един буфер.

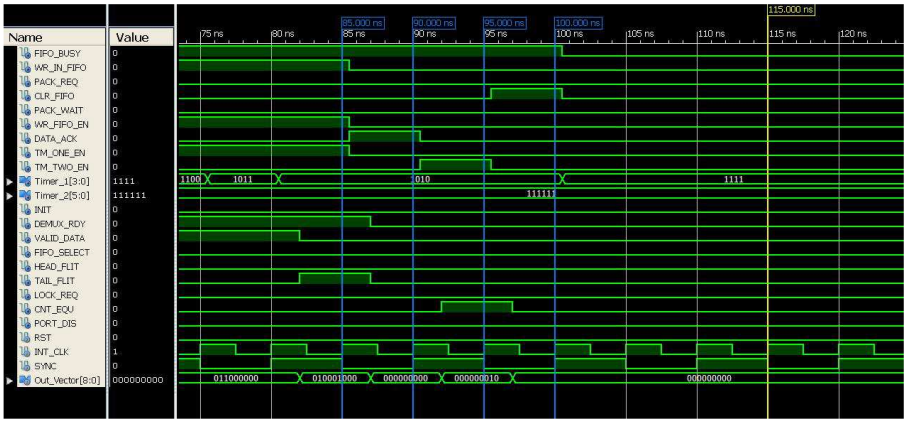
4.3. Експериментални резултати за една FIFO опашка на входен буфер: На Фиг.4.13 е показана генерираната от симулатора времедиаграма за



Фиг.4.13. Времедиаграма за начало на запис на пакет в опашка на входен буфер

началото на запис на пакет в селектирана FIFO опашка на входен буфер. В конкретния случай, след 40-та ns опашката е в състояние *Ready* и очаква пакет. В 47-ма ns сигналът HEAD_FLIT се активира, което показва наличие на заглавен флит от пристигащ пакет, в следствие на което в 50-та ns се преминава в състояние *Write_Packet* (запис на пакет) и веднага се подава заявка PACK_REQ към арбитъра за разпределение на пакета към пожелан изходен канал. В 55-та ns безусловно се преминава в състояние *Wr&Rd_Packet* (запис и четене на пакет), за да може пакетът да се предава към следващия маршрутизатор. Запуска се Timer_1 за изход по Time Out. В 60-та ns се вижда първата промяна състоянието на Timer_1.

На Фиг.4.14 е показана генерираната от симулатора времедиаграма на краят на запис на пакет в селектирана FIFO опашка на входен буфер. В конкретния случай, индицирането за край на записа става в 82-ра ns, когато сигналът TAIL_FLIT се активира и показва наличие на приет опашен флит. В 85-та ns опашката преминава в състояние *Read_Packet* (четене на пакет), Timer_1 спира, а в 90-та ns безусловно се преминава в състояние *End_Packet* (край на четене на пакет), в което се очаква окончателно прочитане на пакета от опашката или изход по TimeOut по Timer_2. В случая в 92-ра ns условието (CNT_EQU=1) индицира край на четене от тази опашка, в 95-та ns се преминава в състояние *Init* (инициализация), а в 100-на ns се преминава в *Idle* (неактивно), където опашката чака да бъде селектирана отново от входния буфер за запис на друг пакет.

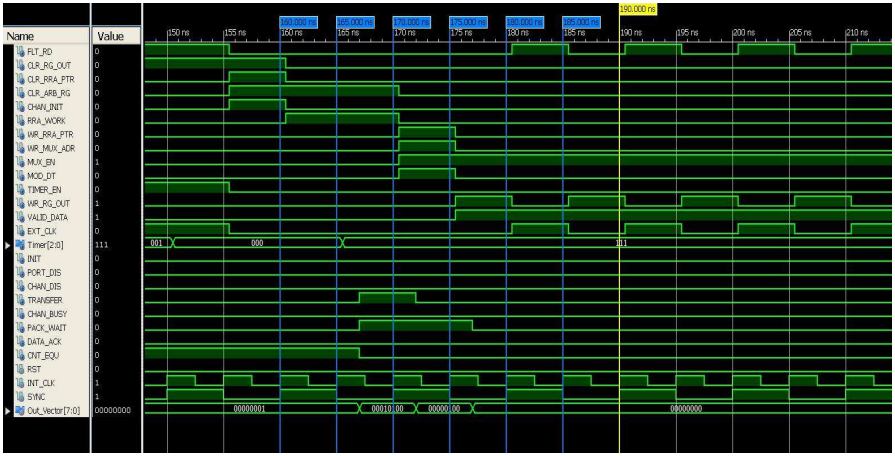


Фиг.4.14. Времедиаграма за край на запис на пакет в опашка на входен буфер

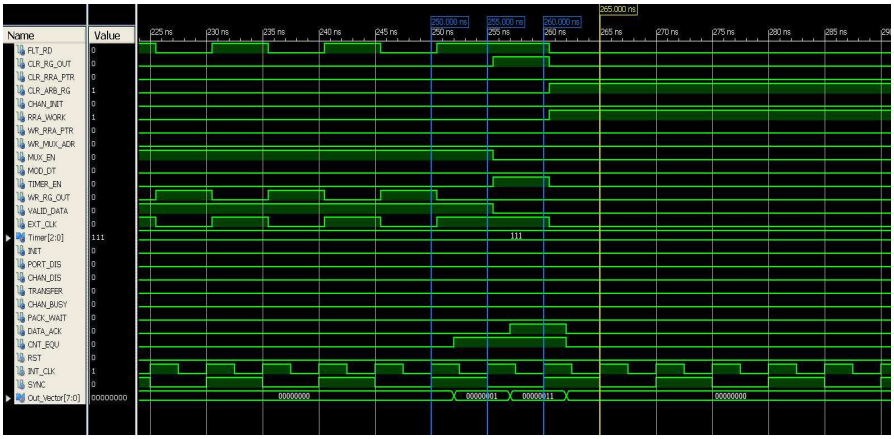
4.4. Експериментални резултати за RRA и изходен канал: На Фиг.4.21 е показана генерираната от симулатора времедиаграма на начало на предаване на пакет от изходен канал на текущия маршрутизатор към входен канал на съседен маршрутизатор. В конкретния случай в 160-та ns започва фазата на разпределение и арбитраж на входните опашки с пакети, кандидати за текущия изходен канал. Сигналът TRANSFER (показващ вече осъществено разпределение) се активира в 167-ма ns, а в 170-та ns изходният канал преминава в състояние *Path_Setting*, установява се път на пакета през комутатора

(WR_MUX_ADR=1), преинициализира се кръговия указател за приоритет на RRA арбитъра (WR_RRA_PTR=1) и се разрешава модификацията на заглавния флит на пакета (MOD_DT=1). Готовността на входния канал на съседния маршрутизатор, декларирана чрез (PACK_WAIT=1)∩CHAN_BUSY=0) предизвиква преминаване на изходния канал в състояние *Transfer_Begin* (начало на предаване) в 175-та ns и безусловен преход в състояние *Transfer_1* в 180-та ns. Същинското предаване започва в състоянието *Transfer_Begin* в 175-та ns с активирането на сигнала VALID_DATA и записа на първия флит от пакета в регистъра на изходния канал по предния фронт на сигнала WR_RG_OUT.

На Фиг.4.22 е показана генерираната от симулатора времедиаграма за край



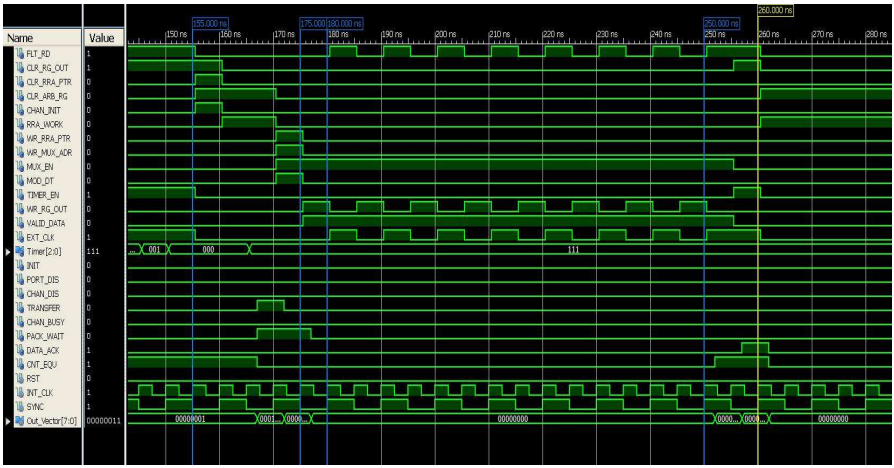
Фиг.4.21. Времедиаграма за начало на предаване на пакет от изходен канал към входен буфер на съседен маршрутизатор



Фиг.4.22. Времедиаграма за край на предаване на пакет от изходен канал към входен буфер на съседен маршрутизатор с потвърждение на успешен трансфер

на един успешен трансфер на пакет от изходен канал на текущия маршрутизатор към входен канал на съседен маршрутизатор. В конкретния случай в 252-та ns се активира сигналът CNT_EQU=1 (пакетът от опашката е прочетен), който в 255-та ns предизвиква преход на изходния канал в състояние *Transfer_End* (край на трансфера). Данните по линиите за връзка стават невалидни (VALID_DATA=0) и се стартира Timer (TIMER_EN=1) за изход по Time Out. В 257-ма ns сигналът за потвърждение на успешно приет пакет DATA_ ACK от съседния маршрутизатор се активира и в 260-та ns по условие (CHAN_DIS=0)∩ (PORT_DIS=0)∩(INIT=0)∩ (SYNC=1) изходния канал преминава в състояние *Routing_and_Arbitrage_1* за участие в следващо разпределение.

На Фиг.4.23 е показана генерираната от симулатора времедиаграма за цялостния трансфер на същия пакет, чиито начало и край на обмен са показани съответно на Фиг.4.21 и Фиг.4.22. Пакетът е с дължина осем флита и неговото предаване се осъществява при благоприятни за това условия (свободен буфер в съседния маршрутизатор и успешен арбитраж за един такт). Времедиаграмата показва каква е минималната латентност на този пакет в текущия маршрутизатор. В конкретния случай латентността е два такта (в интервала между 155-та ns и 175-та ns при условие, че вътрешния тактов сигнал INT_CLK е с период 5ns) и включва следните действия: 1). Запис на заглавния флит на пакета в селектирана опашка на входния канал; 2) Маршрутизация и арбитраж (разпределение на пакета към изходен канал); 3) Формиране на път за пакета през комутатора към изходния канал. Трябва да бъде подчертано, че латентността не зависи от дължината на пакета и като пряко следствие – използваемостта на един линк е толкова по-голяма, колкото по-дълги са пакетите, които се предават през него.



Фиг.4.23. Времедиаграма за трансфер на пакет с дължина осем флита от текущия към съседния маршрутизатор

Основни изводи и предложения за практиката

Един много важен проблем при системните мрежи на паралелните компютри

и маршрутизаторите, които са основните модули в тях е, че те могат съществено да намалят пропускателната способност на комуникационната подсистема на един паралелен компютър, което оказва съществено влияние върху неговата производителност. Осигуряването на мащабируемост, толерантност към откази и възможност за реконфигурация допълват ключовите изисквания към мрежата, които са взети предвид при създаването на архитектурната платформа на маршрутизатора, адаптиран към DLH мрежова топология. В процеса на проектирането и изследването на тази платформа са направени следните основни изводи:

Разработеният алгоритъм за минимална адаптивна маршрутизация на пакети, адаптиран към комуникационна мрежа с DLH топология, и проектираният блок за модификация на заглавния флит, осигуряващ работата на алгоритъма, позволяват действията за маршрутизацията и арбитража на преминаващите през маршрутизатора пакети да бъдат извършени само за един такт.

Предложени са архитектура на входен буфер на маршрутизатор на базата на пул от FIFO опашки и архитектура на една FIFO опашка с техните компоненти, връзки и управление. При тази архитектура се осигурява пряко свързване на опашките към изходните канали на маршрутизатора, което позволява в един и същи момент от време свързването на повече от една опашка към изходните канали на маршрутизатора.

Разработен е блок за маршрутизация и арбитраж с арбитри при всеки от изходните канали, който реализира разпределение на входните опашки с пакети към изходите с една итерация за един такт. Проектираният за него арбитър на входна FIFO опашка, работещ в условията на текущото локално състояние на мрежата, също стои в основата на реализирането на адаптивната маршрутизация. Кръговите арбитри при изходните канали осигуряват високата скорост и справедливост при обслужване на заявките.

Разработени са алгоритми за работата на входните буфери с техните опашки, изходните канали, комутатора и арбитрите. Проектирани са техните управляващи автомати (състояния, преходи, входни и изходни сигнали и взаимодействия), реализиращи паралелната синхронна работа на тези блокове в маршрутизатора.

Създадени са програми, описващи работата на управляващите блокове на маршрутизатора в съответствие с техните алгоритми. Чрез тези програми е симулирана работата при управлението на маршрутизатора. На тази база е изследвано поведението и са визуализирани действията на основните модули в зависимост от текущите им състояния и оказаните входни въздействия върху тях.

Експерименталните резултати от изследванията доказват правилната работа на функционалните възли съгласно поставените изисквания. Измерванията, визуализирани върху времедиаграмите, са анализирани в съответствие със зададените алгоритми за работа (състояния и преходи) и чрез броя тактове, за които се извършва управление на ресурсите и преноса на данните. Получените резултати доказват постигане на висока скорост и минимална латентност от два такта при трансфера на пакетите през маршрутизатора.

Приноси по дисертационния труд

Приноси с научно-приложен характер:

1. Предложена е концептуална платформа за проектиране на високоскоростен маршрутизатор за паралелни компютри, изградени на базата на DLH мрежова топология, осигуряващ ниска латентност, висока отказоустойчивост и липса на взаимна блокировка.
2. Разработен е алгоритъм за минимална адаптивна маршрутизация на пакети за комуникационна мрежа с DLH топология, подходящ за реализация чрез апаратни средства, чрез който само за един такт се определя посоката на предаване на един пакет.
3. Предложена е архитектура на високоскоростен маршрутизатор, който се характеризира с тристепенен конвейер, пряко свързване на входните опашки към комутатор, арбитраж при всеки от изходните канали и Cut-Through техника за комутация.
4. Предложена е архитектура и е разработен алгоритъм на функциониране на входен буфер на маршрутизатора на базата на пул от FIFO опашки, директно свързани към комутационните канали и позволяващи в даден интервал от време от един и същ входен буфер да бъдат изпращани към изходните канали множество пакети за осигуряване на максимална пропускателна способност. Подходът позволява достигане на пълно натоварване на изходните канали.

Приноси с приложен характер:

5. Разработен е блок за маршрутизация и арбитраж – двустепенен разпределител с iSLIP алгоритъм, изграден от арбитри за всяка от опашките на входните буфери и за всеки от изходните канали. Разпределението се извършва с една итерация във всеки времеви интервал.
6. Разработени са структурите и алгоритмите на функциониране на основните възли на един маршрутизатор за DLH мрежова топология – буфери, опашки, арбитри и канали, в тяхната цялост и взаимосвързаност в маршрутизатора.
7. На основата на симулационни експерименти в среда Verilog при подходящо подбрани входни вектори, отразяващи критични за производителността на маршрутизатора ситуации, е доказана работоспособността на основните функционални възли на маршрутизатора и минималната латентност от два такта при трансфера на един пакет.

Списък на публикуваните работи по темата на дисертацията

1. Angelov M. Routers for MPP Computers, Using Direct Communications Networks, John Atanasoff Society of Automatics and Informatics, International Conference AUTOMATICS AND INFORMATICS'2014 October 1-3, 2014, Sofia, Bulgaria, ISSN 1313-1869.
2. Angelov M., Ruskova N. Packet Transfer in DLH Networks, IOSR Journal of Computer Engineering (IOSR-JCE), e-ISSN: 2278-0661, p-ISSN: 2278-8727, Vol.19, Issue 5, Ver. II (Sep.- Oct. 2017), pp. 08-13.
3. Ангелов М. Арбитраж на пакети в маршрутизатор за MPP компютри с DLH мрежова топология, Трета научна конференция с международно участие "Компютърни науки и технологии", 25-26 Септември, 2015, Варна, България, ISSN 1312-3335, Бр.1/2015, стр. 38-45.
4. Ангелов М. Един вариант за Cut-Trough управление на потока в маршрутизатор за MPP компютри, Трета научна конференция с международно участие "Компютърни науки и технологии", 25-26 Септември, 2015, Варна, България, ISSN 1312-3335, Бр.1/2015, стр. 46-54.
5. Ангелов М. Маршрутизация на пакети в MPP компютри с DLH мрежова топология, Втора научна конференция с международно участие "Компютърни науки и технологии", 26-27 Септември, 2014, Варна, България, ISSN 1312-3335, Бр.1/2014, стр. 64-69.
6. Ангелов М. Структура и управление на буфер за входен канал на маршрутизатор за MPP компютри, Втора научна конференция с международно участие "Компютърни науки и технологии", 26-27 Септември, 2014, Варна, България, ISSN 1312-3335, Бр.1/2014, стр. 71-76.

ABSTRACT

of the main part of a dissertation on the subject: Router Architecture for MPP and NUMA Computers with DLH Network Topologies

of the Requirement for the Degree Doctor of Philosophy

by Mr Milen Georgiev Angelov

Assuring maximal performance in modern computing systems is done by utilizing various types of parallelism at all architectural levels. Communication subsystems, with which the the processing elements exchange information, have a significant impact of performance. The communication subsystem is so important to a parallel system that many of its performance characteristics are expressed in terms of the inter-processor communication times. This emphasizes that communication problems are particularly relevant. Therefore, increasing the performance of a parallel computer requires solving the problem of synthesizing an effective communication subsystem.

This dissertation is on research in creating an architectural platform for a high-speed router for use in a parallel computer, which uses a double-loop hypercube (DLH) (a modified hypercube) network topology for its communication subsystem. The designed router has high throughput, low latency, no deadlocking and a high degree of failure tolerance.

The first chapter of a dissertation presents a relevant literature survey regarding the aforementioned issues. It presents and analyzes parallel computers and their fundamental characteristics, primarily with regard to memory organization. A more in-depth analysis is presented on the topologies used to connect the nodes of a parallel system and the communication networks which carry out the exchange of data. The analysis and these comparisons lead to some relevant conclusions and allow the primary purpose of the research to be formulated.

The second chapter presents a comparative analysis of structures, methods and algorithms used in the design of routers for system networks in parallel computers. These include types of routing, techniques for packet switching and flow control, fundamental architectural models, functional blocks for creating data-flow paths, input buffers, switches and output channels, and control blocks: arbiters, allocators and link controllers. This motivates the choices made with regard to operating principles, structures and algorithms, which form the basis of the project. These are the solutions to some of the issues to achieving the main goal of the research.

The third chapter presents a series of mutually justified architectural and technical solutions describing the design of an architectural platform of a high speed router for a parallel computer with a DLH topology. To this end, the main tasks which were carried out were:

- the formulation of the fundamental principles, requirements and constraints needed to outline the creation of this architectural platform;

- the development of an algorithm for minimal adaptive routing of packets, adapted to communication networks with DLH topologies, including a block for modifying the so-called mismatch field, which the routing algorithms uses;

- the development of a block diagram of a high-speed router based on the following fundamental principles: 1) a three-step pipeline for flit transfer; 2) direct

connections from the queues of the input buffers to the switch; 3) arbitration at every output channel; 4) utilization of cut-through switching techniques;

- the development of the router's input buffer – a pool of FIFO queues and their control, including the queue's architecture (components, connections and control logic);

- the development of a block for routing and arbitration made up of arbiters for each input-buffer queue and each output channel, including packet allocation from the input FIFO queues to the outputs which takes one iteration at each time interval, and direct connections between the input queues to the outputs, which together guarantee maximal utilization of the output channels and achieve a high-throughput capability in the router;

- the development of the algorithms for synchronous operation of the input buffers and their queues, the output channels and achieve a high-throughput capability in the router;

The fourth chapter presents results from experiments related to the operation of the router. These consist of simulations done with Xilinx PlanAhead v14.6, written in Verilog. The simulations give insight into and visualize via timing diagrams the operation of the separate functional nodes, as well as their interconnectivity as a single system. Certain conclusions are drawn on the basis of these results, which prove that the functional blocks of the router operate correctly and that low latency is achieved in the transfer of packets through the nodes of the DLH network.