Công trình được hoàn thành tại Khoa Điện tử - Viễn thông, Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội.

Người hướng dẫn khoa học:

1. PGS.TS. Trần Xuân Tú
2. PGS.TS. Ngô Diên Tập

Phản biện 1: PGS.TS Hoàng Trang

Phản biện 2: TS. Nguyễn Ngọc Minh

Phản biện 3: TS. Nguyễn Vũ Thắng

Luận án sẽ được bảo vệ trước Hội đồng cấp nhà nước chấm luận án tiến sĩ họp tại: Phòng 212, Nhà E3, Trường Đại học Công nghệ - ĐHQG Hà Nội vào hồi 13 giờ 30 ngày 21 tháng 12 năm 2017.

Có thể tìm hiểu luận án tại:

- Thư viện Quốc gia Việt Nam
- Trung tâm Thông tin - Thư viện, Đại học Quốc gia Hà Nội.

LỜI CAM ĐOAN

Tác giả xin cam đoan toàn bộ nội dung trình bày trong luận án và các công trình nghiên cứu là của riêng tác giả và nhóm cán bộ hướng dẫn gồm PGS.TS. Trần Xuân Tú và PGS.TS. Ngô Diên Tập. Các số liệu và kết quả trình bày trong luận án là trung thực và chưa từng được ai công bố trong bất kỳ công trình nào trước đó.

Lê Văn Thanh Vũ

LỜI CẢM ƠN

Để hoàn thành luận án này tôi xin gửi lời cảm ơn chân thành đến PGS.TS. Trần Xuân Tú - người rất tận tình, tận tâm giúp đỡ và tạo mọi điều kiện thuận lợi nhất cho tôi trong suốt quá trình thực hiện luận án. Tôi cũng xin bày tỏ lòng biết ơn sâu sắc trước những giúp đỡ và đóng góp ý kiến quý báu của PGS.TS. Ngô Diên Tập để tôi hoàn thành được luận án này.

Tôi xin chân thành cảm ơn các thầy cô, các cán bộ của Phòng thí nghiệm trọng điểm Hệ thống tích hợp thông minh, Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội đã tạo điều kiện thuận lợi và hỗ trợ tôi trong quá trình thực hiện luận án. Xin chân thành cảm ơn các thầy cô ở Khoa Điện tử Viễn thông, Trường Đại học Công nghệ, đặc biệt là các thầy cô ở Bộ môn Điện tử và Kỹ thuật Máy tính đã giảng dạy và tạo mọi điều kiện thuận lợi cho tôi trong suốt thời gian học tập và nghiên cứu tại Nhà trường.

Tôi cũng chân thành cảm ơn sự hỗ trợ cả vật chất và tinh thần của Trường Đại học Khoa học – Đại học Huế; Khoa Điện tử – Viễn thông của Trường Đại học Khoa học – Đại học Huế.

Cuối cùng, tôi xin cảm ơn đến Vợ và Con - Phan Thị Hạnh Nguyên và Lê Vũ Bảo Ngọc là nguồn động lực vô tận để tôi vững bước trên con đường nghiên cứu và thực hiện luận án. Đồng thời, tôi cũng muốn gửi lời cảm ơn đến Cha Mẹ và các anh chị em đã động viên giúp đỡ tôi trong thời gian qua và cả trong tương lai sắp tới.

Mục lục

Mục lục **i**

Danh mục ký hiệu và chữ viết tắt **iv**

Danh mục các bảng **vii**

Danh mục các hình vẽ, đồ thị **viii**

Mở đầu **1**

Chương 1: Tổng quan về mạng trên chip **5**

1.1 Giải pháp truyền thông mạng trên chip 6

1.2 Cấu trúc liên kết 9

1.3 Kỹ thuật truyền thông 12

1.3.1 Cơ chế điều khiển luồng 15

1.3.2 Cơ chế điều chuyển dữ liệu 17

1.3.3 Chiến lược bộ đệm 18

1.4 Giải thuật định tuyến 21

1.4.1 Phân loại định tuyến 22

1.4.2 Giải thuật định tuyến tĩnh 23

1.4.3 Giải thuật định tuyến thích nghi 25

1.4.4 Thực hiện định tuyến 29

1.5 Kết luận chương 31

Chương 2: Vấn đề tái cấu hình và truyền thông tái cấu hình **33**

2.1 Hệ thống trên chip và định hướng tái cấu hình 34

2.2 Vấn đề tái cấu hình đối với mạng trên chip 38

2.2.1 Động lực và thách thức của mạng trên chip tái cấu hình . 39

2.2.2 Phân loại giải pháp tái cấu hình mạng trên chip 42

2.3 Một số kiến trúc mạng trên chip tái cấu hình điển hình 46

2.3.1 Tái cấu hình cấu trúc liên kết 47
2.3.2 Tái cấu hình kiến trúc bộ định tuyến 48
2.4 Các vấn đề cần quan tâm khi xây dựng giải pháp truyền thông tái cấu hình 52
2.4.1 Hoạt động quản lý cấu hình và điều khiển truyền thông . 53
2.4.2 Giải thuật định tuyến cho mạng trên chip tái cấu hình . . 55
2.5 Tổng kết chương 58
Chương 3: Giải pháp tái cấu hình cho mạng trên chip 63
3.1 Cơ sở thực hiện giải pháp tái cấu hình 64
3.1.1 Một số định nghĩa 64
3.1.2 Cơ sở giải pháp tái cấu hình 66
3.2 Giải pháp cập nhật thông tin định tuyến 70
3.2.1 Cập nhật định tuyến khi bộ định tuyến bị cấm nằm trên đoạn thẳng định tuyến 70
3.2.2 Cập nhật định tuyến khi bộ định tuyến bị cấm nằm tại góc định tuyến 72
3.2.3 Cập nhật định tuyến khi bộ định tuyến bị cấm nằm ở lân cận góc định tuyến 73
3.3 Kiến trúc bộ định tuyến tái cấu hình đề xuất 75
3.3.1 Giải pháp kiến trúc cho bộ định tuyến 75
3.3.2 Kiến trúc chi tiết khối cổng lối vào và lối ra 79
3.3.3 Kiến trúc thực hiện hoạt động cập nhật định tuyến 82
3.4 Mô hình hóa, kiểm chứng và thực thi 86
3.4.1 Mô hình hóa kiến trúc bộ định tuyến tái cấu hình 87
3.4.2 Thực thi kiến trúc 90
3.5 Kết luận chương 94
Chương 4: Đánh giá hiệu năng truyền thông và hiệu quả giải pháp tái cấu hình 96
4.1 Hoạt động đánh giá hiệu năng truyền thông mạng trên chip . . . 97
4.1.1 Thông số đánh giá hiệu năng truyền thông 97
4.1.2 Phương pháp đánh giá 99
4.2 Đánh giá hoạt động truyền thông trên chip 101
4.2.1 Thiết kế cho đánh giá truyền thông 101

4.2.2 Kịch bản đánh giá 102
4.2.3 Kết quả đánh giá 104
4.3 Đánh giá giải pháp tái cấu hình mạng trên chip 107
4.3.1 Phương pháp đánh giá giải pháp tái cấu hình mạng trên chip 107
4.3.2 Mô phỏng và kết quả 109
4.4 Kết luận chương 115
Kết luận và hướng phát triển **117**
Danh mục công trình khoa học của tác giả liên quan đến luận án **120**
Tài liệu tham khảo **121**

Danh mục các ký hiệu và chữ viết tắt

Từ viết tắt	**Từ tiếng Anh**	**Mô tả**
μPU	micro Processor Unit	Đơn vị vi xử lý
ADC	Analog Digital Converter	Bộ chuyển đổi tương tự - số
ALU	Arithmetic Logic Unit	Đơn vị tính toán số học
ASIC	Application Specific Integrated Circuit	Mạch tích hợp chuyên dụng
ASIP	Application Specific Instruction Processor	Vi xử lý với tập lệnh chuyên dụng
BIST	Built-In Self-Test	Thiết kế có cơ chế tự kiểm tra ngay bên trong
CEB	Configuration Exchange Buffer	Bộ đệm trao đổi cấu hình
CF	Congestion Factor	Hệ số tắc nghẽn
CGRA	Coarse-Grained Reconfigurable Architecture	Kiến trúc tái cấu hình hạt nhân thô
Complex System	Complex System	Hệ thống gồm nhiều lõi chức năng khác nhau cùng tích hợp trên một chip tạo nên một vi mạch hoàn thiện
DAC	Digital Analog Converter	Bộ chuyển đổi số-tương tự
DEB	Data Exchange Buffer	Bộ đệm trao đổi dữ liệu
DF	Distance Factor	Hệ số khoảng cách
DFT	Design for Test	Thiết kế cho phép tự kiểm tra sai hỏng
DOR	Dimenssion Order Routing	Giải thuật định tuyến ưu tiên theo hướng
DSM	Deep Submicron	Công nghệ chế tạo nhỏ hơn micro

DSP	Digital Signal Processor	Bộ xử lý tín hiệu số
flit	Flow control unit	Đơn vị thông tin cơ bản trong truyền thông trên chip
FPGA	Field-Programmable Gate Array	Mảng cổng logic khả trình
FGRA	Fine-Grained Reconfigurable Architecture	Kiến trúc tái cấu hình hạt nhân tinh
FSM	Finite State Machine	Máy trạng thái hữu hạn
GPP	General Purpose Processor	Bộ vi xử lý đa dụng
HOL	Head-of-line block	Khóa đầu tuyến
HRE	Heterogeneous Reconfigurable Engine	Công cụ tái cấu hình không đồng nhất
IPcore	Intellectual Property core	Lõi IP
ISRC	Intelligent Service Reconfigure Computing	Công cụ tính toán tái cấu hình thông minh
LUT	Look-Up Table	Bảng tìm kiếm
MGRA	Middle-Grained Reconfigurable Architecture	Kiến trúc tái cấu hình hạt nhân trung
NF algorithm	Negative First algorithm	Giải thuật định tuyến phía Âm trước
NL algorithm	North Last algorithm	Giải thuật định tuyến phía Bắc cuối
NI	Network Interface	Giao tiếp mạng
NoC	Network-on-Chip	Mạng trên chip
NRE cost	Non Recurring Engineering cost	Chi phí nghiên cứu lần đầu
OCP	Open Core Protocol	Chuẩn giao tiếp lõi mở
OE routing	Odd-Even Routing	Giải thuật định tuyến chẵn lẻ
Pflit	Physical flow control unit	Đơn vị thông tin vật lý lan truyền qua các liên kết của các bộ định tuyến
PTT	Path-To-Target	Trường định tuyến
QoS	Quality of Service	Đảm bảo chất lượng dịch vụ
	Router	Bộ định tuyến

RAM	Random Access Memory	Bộ nhớ truy xuất ngẫu nhiên
RNoC	Reconfigurable Network-on-Chip	Mạng trên chip có khả năng tái cấu hình
RPU	Reconfigurable Processing Unit	Đơn vị xử lý tái cấu hình
RSoC	Reconfigurable System-on-Chip	Hệ thống trên chip có khả năng tái cấu hình
RTL	Register Transfer Level	Mức chuyển dịch thanh ghi
SoC	System-on-Chip	Hệ thống trên chip
SAF	Store and Forward	Cơ chế lưu và chuyển tiếp
SXY routing	Surround XY routing	Giải thuật định tuyến đi vòng XY
Topology	Topology	Cấu trúc liên kết mạng
TLM	Transaction-Level Modeling	Mô hình hóa mức giao dịch
VCPIQ	Virtual Channel Priority Input Queuing	Hàng đợi lối vào ưu tiên kênh ảo
VCT	Virtual Cut Through	Cơ chế xuyên cắt ảo
VOQ	Virtual Output Queuing	Hàng đợi lối ra ảo
WF algorithm	West First algorithm	Giải thuật định tuyến phía Tây trước
WH	Wormhole	Cơ chế Wormhole

Danh sách bảng

1.1 So sánh ưu nhược điểm của bus và mạng trên chip [25] 7

2.1 Bảng tóm tắt các công trình mạng trên chip tái cấu hình. 60

3.1 Bảng các trạng thái hoạt động của cổng vào và cổng ra. 90

3.2 Bảng so sánh kết quả thực thi bộ định tuyến của RNoC và một số công trình khác. 93

4.1 Thông số của mạng trên chip trong mô phỏng đánh giá 104

4.2 Thông số của mạng trên chip tái cấu hình trong mô phỏng đánh giá109

4.3 Tài nguyên truyền thông trong mỗi bộ định tuyến 114

Danh sách hình vẽ

1.1 Mô hình mạng trên chip cơ bản. 6
1.2 Một số dạng cấu trúc liên kết mạng trên chip. 10
1.3 Cấu trúc của một bộ định tuyến dùng trong mạng 2D-mesh [34]. 11
1.4 Kỹ thuật truyền thông cho mạng trên chip. 12
1.5 Mô hình phân lớp áp dụng cho mạng trên chip. 13
1.6 Các cơ chế điều chuyển dữ liệu. 17
1.7 Chiến lược bộ đệm lối vào và lối ra. 19
1.8 Chiến lược bộ đệm lối ra ảo. 20
1.9 Chiến lược bộ đệm lối vào ưu tiên kênh ảo. 21
1.10 Hiện tượng khóa vòng. 22
1.11 Lưu đồ thuật toán định tuyến XY. 24
1.12 Định tuyến tĩnh XY và YX trong mạng 2D-mesh. 25
1.13 Góc mở trong giải thuật định tuyến góc mở. 26
1.14 Lưu đồ thuật toán định tuyến Phía Tây trước (WF). 27
1.15 Các đường định tuyến của nhóm giải thuật Turn model. 27
1.16 Lưu đồ thực hiện giải thuật định tuyến chẵn-lẻ [10]. 29

2.1 Mạch cộng tái cấu hình bằng FPGA. 34
2.2 Kiến trúc cơ bản hệ thống tái cấu hình hạt nhân thô. 37
2.3 Hoạt động tái cấu hình dành cho mạng trên chip. 44
2.4 Kiến trúc mạng và bộ định tuyến của ReNoC [61]. 47
2.5 Hoạt động thiết lập cấu trúc liên kết ảo [71]. 49
2.6 Kiến trúc bộ đệm đa năng trong ViChar [50]. 50
2.7 Kiến trúc bộ định tuyến chia sẻ tài nguyên bộ đệm [11]. 51
2.8 Kiến trúc bộ định tuyến liên kết hai chiều (BiNoC) [33]. 52

2.9 Liên kết nối giữa các bộ định tuyến trong dạng lưới hai chiều 2D-mesh [75]. 53
2.10 Giải pháp dịch chuyển tác vụ truyền thông [51]. 54
2.11 Giải pháp thu hồi tài nguyên của ROAdNoC [17]. 55
2.12 Giải pháp định tuyến phân tán trên cơ sở logic (LBDR) [19]. . . . 56
2.13 Nguyên lý kiến trúc bộ định tuyến kết hợp DyAD. 57
2.14 Giải thuật định tuyến vòng trong DyNoC [7]. 58

3.1 Mạng trên chip dạng lưới hai chiều kích thước 5×5 với thuật toán định tuyến XY. 65
3.2 Bộ định tuyến căn bản cho mạng trên chip 2D-mesh. 67
3.3 Cấu trúc gói tin. 68
3.4 Cập nhật định tuyến khi bộ định tuyến bị cấm nằm trên đoạn thẳng định tuyến (trường hợp A). 71
3.5 Thông tin trường định tuyến tương ứng với Hình 3.4. 71
3.6 Cập nhật định tuyến khi bộ định tuyến bị cấm nằm tại góc định tuyến (trường hợp B). 72
3.7 Thông tin trường định tuyến tương ứng với Hình 3.6. 73
3.8 Cập nhật định tuyến khi bộ định tuyến bị cấm ở lân cận góc định tuyến (trường hợp C). 74
3.9 Thông tin trường định tuyến tương ứng với Hình 3.8a. 74
3.10 Thông tin trường định tuyến tương ứng với Hình 3.8b. 75
3.11 Kiến trúc bộ định tuyến có thể tái cấu hình đề xuất. 77
3.12 Cơ chế chuyển trạng thái của bộ định tuyến RNoC. 78
3.13 Kiến trúc chi tiết của khối cổng lối vào. 80
3.14 Lưu đồ hoạt động của khối cổng lối vào. 81
3.15 Kiến trúc chi tiết của khối cổng lối ra. 83
3.16 Lưu đồ hoạt động của khối cổng lối ra. 84
3.17 Kiến trúc chi tiết của khối cổng ảo thực thi tái cấu hình. 85
3.18 Quá trình thực hiện giải pháp mạng trên chip tái cấu hình. 86
3.19 Chi tiết kiến trúc bộ định tuyến cho RNoC. 87
3.20 Sơ đồ trạng thái của máy trạng thái hữu hạn cho khối cổng lối vào. 88
3.21 Sơ đồ trạng thái của máy trạng thái hữu hạn cho khối cổng lối ra. 89

3.22 Quy trình nghiên cứu, thiết kế và thực hiện vi mạch số. 92

4.1 Nền tảng đánh giá truyền thông mạng trên chip. 102
4.2 Lược đồ hoạt động phát tải vào mạng. 103
4.3 Độ trễ truyền và thông lượng theo hệ số tải với gói dữ liệu 16*flit*. 105
4.4 Trễ truyền thông đánh giá theo hệ số tải và kích thước gói tin. . . 106
4.5 Thông lượng mạng theo hệ số tải và kích thước gói tin. 106
4.6 Trễ truyền tương ứng với các vị trí bị cấm. 110
4.7 Thông lượng mạng tương ứng với các vị trí bị cấm. 111
4.8 So sánh giá trị thông lượng của các giải pháp tái cấu hình. 113
4.9 Độ trễ truyền của mạng trên chip tái cấu hình ở mô hình mức RTL.114
4.10 So sánh tỷ lệ thông lượng với kích thước bộ đệm và chi phí không gian thực thi của các kiến trúc tái cấu hình. 115

Mở đầu

Mô hình truyền thông trên chip phổ biến hiện này được xây dựng dựa vào kết nối điểm-điểm, kiến trúc bus truyền thống (hoặc kiến trúc bus phân tầng). Với kiến trúc này, một bộ phân xử bus sẽ đóng vai trò cấp phát quyền truy cập bus cho các thành phần truyền thông nhằm tránh các xung đột trong quá trình trao đổi thông tin trên bus. Kiến trúc truyền thông bus có một số hạn chế căn bản như: băng thông bị giới hạn, khả năng mở rộng kém... Các hệ thống phức hợp đòi hỏi nhu cầu truyền thông cao, việc xây dựng một mô hình truyền thông mới và hiệu quả là hết sức cần thiết.

Mạng trên chip (NoC: *Network-on-Chip*) được đề xuất là một giải pháp toàn diện cho sự phát triển của các hệ thống trên chip phức hợp trong xu thế thiết kế lấy truyền thông làm trung tâm. Kiến trúc truyền thông mạng trên chip thực hiện các giao tác truyền thông bằng nguyên lý phân đoạn trên một cấu trúc liên kết gồm nhiều bộ định tuyến, kết hợp với các kỹ thuật truyền thông phù hợp để nâng cao hiệu quả hoạt động của toàn hệ thống. Mô hình mạng trên chip cho phép người thiết kế tích hợp ngày càng nhiều lõi IP trong một hệ thống nhằm đáp ứng yêu cầu ngày càng cao của các ứng dụng. Điều này cũng tạo nên áp lực thúc đẩy các nghiên cứu nhằm hoàn thiện mô hình mạng trên chip. Trong đó, định hướng nghiên cứu để nâng cao khả năng linh hoạt trong truyền thông rất được quan tâm nghiên cứu nhằm đáp ứng các yêu cầu phát triển của hệ thống phức hợp đa dạng, tạo nên các mạng trên chip tái cấu hình. Nghiên cứu hoạt động tái cấu hình mạng trên chip đã và đang được phát triển rất đa dạng dựa theo các chức năng truyền thông cụ thể trong mô hình phần lớp; có thể dựa vào khả năng quản trị hệ thống để điều khiển quá trình truyền thông, áp dụng các giải pháp truyền thông tự thích ứng hoặc các kiến trúc bộ định tuyến có khả năng tái cấu hình cho mạng trên chip. Hoạt động truyền thông linh hoạt cho mạng trên chip dựa trên khả năng tự thích ứng của các giải pháp truyền thông hiện đang được chú trọng phát triển với nhiều công trình nổi bật đã được

công bố trong những năm gần đây. Các giải thuật định tuyến tạo nên nguyên lý lựa chọn đường đi của các giao tác truyền thông trên tập các tài nguyên truyền thông của mạng trên chip. Hoạt động cập nhật định tuyến sẽ là giải pháp tối ưu để vừa bảo đảm các giao tác truyền thông tin cậy với khả năng thích ứng với sự thay đổi cấu hình mạng trong giải pháp tái cấu hình và hiệu năng truyền thông của định tuyến tĩnh.

Mục tiêu nghiên cứu của luận án là đề xuất giải pháp tái cấu hình cho các kiến trúc truyền thông mạng trên chip, có thể ứng dụng trong các hệ thống phức hợp, có độ tích hợp cao. Việc xây dựng được các giải pháp tái cấu hình cho hệ thống phức hợp sẽ cho phép người thiết kế xây dựng được các kiến trúc chung không chỉ một ứng dụng mà là một dải ứng dụng khác nhau. Kiến trúc đề xuất này có thể tái cấu hình tùy theo thực trạng của hệ thống, yêu cầu cụ thể của ứng dụng, thậm chí từng phiên bản của một ứng dụng, tạo nên sự mềm dẻo, linh hoạt trong thiết kế.

Đối tượng nghiên cứu ở đây là mô hình mạng trên chip. Để đơn giản hơn, mô hình mạng trên chip với cấu trúc liên kết dạng lưới hai chiều (2D-mesh) được lựa chọn. Tuy nhiên, phương pháp có thể mở rộng cho mô hình mạng ba chiều (3D) hoặc/và các cấu trúc liên kết khác.

Hoạt động nghiên cứu mạng trên chip tái cấu hình là một hướng nghiên cứu thiết kế vi mạch cụ thể kết hợp với quá trình tìm hiểu và xây dựng một giải pháp toàn diện cho mạng tái cấu hình. Do vậy, quá trình thực hiện luận án này sử dụng ba phương pháp nghiên cứu chính gồm:

- Tập hợp tài liệu liên quan và nghiên cứu các vấn đề liên quan đến truyền thông trên chip và mô hình mạng trên chip để làm cơ sở cho việc xây dựng mạng trên chip, và cũng là nền tảng để đi sâu nghiên cứu hoạt động tái cấu hình mạng trên chip.

- Đề xuất các giải pháp truyền thông linh hoạt cho phép mạng có khả năng tự thích ứng với các thay đổi cấu hình.

- Sử dụng phương pháp mô phỏng kết hợp với các công cụ chuyên dùng cho lĩnh vực thiết kế vi mạch (*như: Modelsim, Design Compiler,...*) để mô phỏng và tổng hợp thiết kế.

Giải quyết bài toán truyền thông trên chip các các hệ thống phức hợp cần được xem xét một cách có hệ thống và có trọng tâm để cùng hướng đến giải pháp toàn diện. Xuyên suốt quá trình nghiên cứu thực hiện luận án, các vấn đề

truyền thông mạng trên chip được tổng hợp và sắp xếp để tạo cơ sở lý thuyết. Hướng đến mục tiêu trọng tâm là mạng trên chip tái cấu hình, luận án cũng đã trình bày xu thế tái cấu hình áp dụng cho hệ thống từ mức độ ứng dụng cũng như khả năng tái cấu hình truyền thông thông qua các hoạt động tái cấu hình mạng trên chip. Kết quả đạt được của luận án này tập trung vào ba nội dung chính như sau:

- Đề xuất giải pháp tái cấu hình cho mạng trên chip, cụ thể là: giải thuật cập nhật thông tin định tuyến cho mạng trên chip tái cấu hình có khả năng tự thích ứng với các thay đổi cấu hình do có bộ định tuyến rời khỏi mạng. Giải pháp cập nhật định tuyến cho phép thay đổi đường đi của thông tin linh hoạt để thích ứng với các thay đổi cấu hình mạng trên chip ngay cả khi hệ thống đang hoạt động. Đồng thời giải pháp này vẫn giữ được ưu điểm của hoạt động định tuyến tĩnh tại nguồn cho các giao tác truyền thông không tái cấu hình để bảo đảm được hiệu quả truyền thông tối ưu trên toàn hệ thống. Tiế đó, phát triển kiến trúc bộ định tuyến có khả năng tái cấu hình nhằm thực thi giải pháp nêu trên. Kiến trúc bộ định tuyến đã được mô hình hóa bằng ngôn ngữ phần cứng VHDL ở mức chuyển dịch thanh ghi (RTL: *Register Transfer Level*) và thực thi với công nghệ CMOS 130nm của hãng Global Foundry (Hoa Kỳ).

- Trong luận án này, chúng tôi cũng phát triển phương pháp mô hình hóa, mô phỏng đa lớp sử dụng các ngôn ngữ mô tả phần cứng khác nhau (C++, SystemC, VHDL) để xây dựng nền tảng đánh giá hoạt động truyền thông mạng trên chip. Nền tảng đề xuất cho phép mô phỏng và đánh giá nhanh hiệu năng truyền thông mạng trên chip với các kịch bản đánh giá đa dạng cho cả mạng trên chip và mạng trên chip tái cấu hình.

Luận án được trình bày trong bốn chương với đầy đủ các nội dung liên quan đến hoạt động truyền thông mạng trên chip và định hướng cho giải pháp tái cấu hình. Trong luận án các nội dung trình bày được sắp xếp từ cơ sở lý thuyết cho hoạt động động truyền thông trên chip, sau đó là phần trình bày tổng quan định hướng nghiên cứu tái cấu hình làm cơ sở cho giải pháp được đề xuất. Giải pháp tái cấu hình được thể hiện chi tiết với sự kết hợp từ giải thuật cập nhật định tuyến và hướng đến kiến trúc thực thi của bộ định tuyến. Phương pháp mô phỏng và đánh giá được đưa ra nhằm góp phần khẳng định sự hiệu quả của truyền thông mạng trên chip cũng như giải pháp tái cấu hình đáp ứng cho những yêu cầu truyền thông phức hợp của hệ thống.

Nghiên cứu thiết kế mạng trên chip là một định hướng mới, mà phần lớn tài liệu ở dạng các bài báo khoa học được trình bày theo định hướng nghiên cứu cụ thể. Do đó, luận án này trình bày tổng quan hoạt động truyền thông mạng trên chip một cách ngắn gọn, và có trọng tâm các vấn đề liên quan ở Chương 1. Hoạt động truyền thông mạng trên chip được trình bày có trọng tâm từ góc độ tổ chức mạng, các kỹ thuật truyền thông và các giải thuật định tuyến phù hợp với các ràng buộc của hoạt động thực thi trên vi mạch. Chương 2 trình bày vấn đề tái cấu hình để hướng đến mục tiêu chính của luận án là hoạt động truyền thông mạng trên chip tái cấu hình. Trong chương này vấn đề tái cấu hình mạng trên chip được làm rõ từ góc nhìn tổng quan đến việc sắp xếp phân loại các hướng nghiên cứu đã được công bố.

Giải pháp cập nhật định tuyến cho mạng trên chip tái cấu hình được mô tả chi tiết trong Chương 3 của luận án. Giải pháp cập nhật định tuyến được xây dựng dựa trên đặc điểm hoạt động truyền thông của kiến trúc mạng trên chip với cấu trúc dạng lưới hai chiều 2D-mesh và giải thuật định tuyến tĩnh XY. Để bảo đảm khả năng thích ứng với mọi trường hợp thay đổi cấu trúc mạng giải pháp cập nhật định tuyến chia làm ba trường hợp thay đổi thông tin định tuyến và chúng tôi sử dụng các ngôn ngữ mô tả phần cứng để đề xuất kiến trúc mạng trên chip cho giải pháp tái cấu hình. Chương 4 tập trung cho các hoạt động đánh giá truyền thông của mạng trên chip một cách khách quan theo các chiến lược cụ thể. Từ hoạt động đánh giá truyền thông của mạng trên chip cũng là cơ sở để so sánh hiệu quả của giải pháp tái cấu hình trong nhiều trường hợp khác nhau.

Chương 1

Tổng quan về mạng trên chip

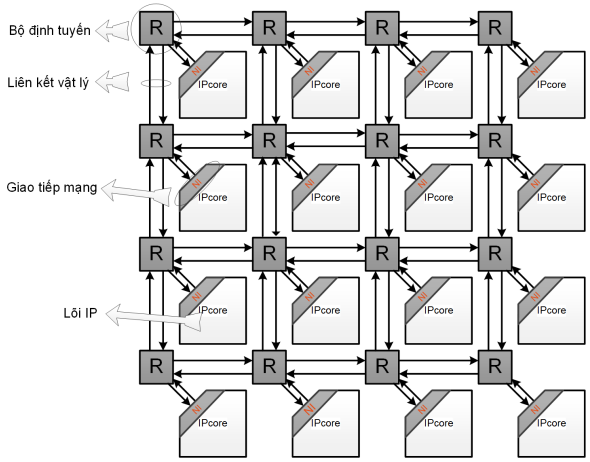
Cùng trong xu thế phát triển của khoa học công nghệ hiện đại, công nghệ điện tử đã, đang và sẽ có nhiều đóng góp sâu và rộng. Các thiết bị điện tử thông minh ngày càng mang lại nhiều lợi ích cho trong nhiều lĩnh vực: y tế, giáo dục, sản xuất kinh doanh...; mà trong đó vi mạch trung tâm tích hợp đa chức năng được gọi là hệ thống trên chip (SoC: *System-on-Chip*). Hệ thống trên chip có thể đơn giản là sự kết hợp của vi xử lý với một số thành phần cơ bản tạo nên một hệ thống điều khiển thiết bị gia dụng (Tivi, máy giặt, ...); hoặc là vi mạch tích hợp đa năng trong các thiết bị điện tử thông minh phổ biến hiện nay. Nhằm đáp ứng ngày các tốt hơn các nhu cầu phát triển, hệ thống trên chip được tích hợp ngày nhiều chức năng hơn, hoạt động phức tạp hơn. Vậy nên, các thiết kế hệ thống trên chip cần được nghiên cứu và giải quyết nhiều thách thức: tích hợp nhiều chức năng, hoạt động đa dạng, yêu cầu tiết giảm năng lượng, tiết kiệm chi phí (cả trong quá trình thiết kế và thực thi trên vi mạch).

Một thách thức lớn của thiết kế các hệ thống phức hợp là khả năng đáp ứng các yêu cầu truyền thông gia tăng nhanh do hệ thống có nhiều chức năng và hoạt động phức tạp. Điều này làm nảy sinh nhiều hạn chế đối với các giải pháp truyền thông trên chip truyền thống (bus chia sẻ, liên kết điểm-điểm). Ngoài ra, hệ thống trên chip còn là sự kết hợp của nhiều lõi IP có hoạt động truyền thông khác biệt (chuẩn giao tiếp, tốc độ, đặc điểm); điều này cần được giải quyết một cách triệt để bằng một cơ chế truyền thông linh hoạt, hướng đến sự cân bằng của chi phí và hiệu quả. Mô hình mạng trên chip (NoC: *Network-on-Chip*) được đề xuất nhằm đáp ứng tốt hơn các yêu cầu truyền thông phức tạp trong các hệ thống trên chip thế hệ mới hiện nay [3]. Nhờ các ưu điểm vượt trội về khả năng mở rộng kiến trúc, tái sử dụng thiết kế và hiệu quả truyền thông cao [25]. Mạng trên chip được xem là giải pháp đáp ứng toàn diện yêu cầu truyền thông các hệ thống phức hợp và là xu thế phát triển của lĩnh vực thiết kế vi mạch

hiện nay. Qua thời gian nghiên cứu phát triển, mô hình mạng trên chip cũng đã dần hoàn thiện và đưa ra một số sản phẩm thương mại nổi trội như dòng sản phẩm FlexNoC của hãng Arteris [37], dòng sản phần CoreLink Interconnect (CMN-600, CCN và CCI) của hãng ARM [39]. Dòng sản phẩm FlexNoC cũng được phát triển trong sản phẩm thương mại ứng dụng vi mạch điều khiển ổ cứng SSD của hãng Arteris [38].

1.1 Giải pháp truyền thông mạng trên chip

Hệ thống trên chip là một hệ thống bao gồm nhiều thành phần chức năng được tích hợp trên một chip đơn. Hệ thống trên chip gồm nhiều lõi chức năng (lõi IP), như là: vi xử lý, bộ chuyển đổi tín hiệu (ADC, DAC), bộ xử lý tín hiệu số (DSP: *Digital Signal Processor*, ...) được liên kết trên một kiến trúc truyền thông phù hợp. Kiến trúc truyền thông của hệ thống trên chip cần đáp ứng đầy đủ mọi yêu cầu trao đổi thông tin giữa các lõi IP; đồng thời góp phần quan trọng để tạo nên một hệ thống đồng bộ, hoạt động hiệu quả và ổn định. Hình 1.1 mô tả một hệ thống gồm nhiều lõi IP giao tiếp với nhau qua kiến trúc truyền thông mạng trên chip.



Hình 1.1: Mô hình mạng trên chip cơ bản.

Theo Hình 1.1, mạng trên chip được hình thành từ các thành phần cơ bản: các lõi IP, kiến trúc truyền thông mạng trên chip và khối giao tiếp giữa mạng và lõi IP (NI: *Network Interface*).

- **Lõi IP** là thành phần thực thi các hoạt động chức năng logic của hệ thống đáp ứng cho yêu cầu từ các ứng dụng cụ thể. Các lõi IP có thể được nghiên cứu và thiết kế đáp ứng tối ưu cho hệ thống, và cũng có thể được tái sử

Bảng 1.1: So sánh ưu nhược điểm của bus và mạng trên chip [25]

Truyền thông bus			**Kiến trúc truyền thông NoC**
Chia sẻ các đường dẫn tín hiệu dài cho nhiều lõi IP.	-	+	Liên kết điểm-điểm với đường dẫn ngắn có khả năng mở rộng.
Giới hạn tần số xung nhịp	-	+	Liên kết ngắn và truyền thông dạng đường ống
Thời gian truy xuất bus chia đều cho số thành phần kết nối bus	-	+	Thời gian truy xuất bộ định tuyến ngắn
Trễ truyền không đổi và rất nhỏ khi đã chiếm bus	+	-	Trễ truyền phụ thuộc vị trí tương đối giữa cặp nguồn-đích
Không gian thực thi (area) rất nhỏ	+	-	Bộ định tuyến cần tối ưu để tiết kiệm không gian thực thi
Xử lý trung tâm do đó giảm hiệu quả truyền thông	-	+	Hoạt động phân tán, độ dài gói tin linh hoạt giúp nâng cao hiệu quả truyền thông
Truyền thông trực tiếp từ nguồn đến đích và thông tin quảng bá cho mọi thành phần	+	-	Khóa vòng là vấn đề hạn chế của NoC và cần có giao thức phù hợp cho thông tin quảng bá
Băng thông giới hạn và chia sẻ cho tất cả các lõi chức năng	-	+	Tổng hợp băng thông và phụ thuộc vào quy mô mạng
Hạn chế hỗ trợ truyền thông cho xử lý song song thời gian thực	-	+	Truyền thông linh hoạt thời gian thực qua mạng
Tiêu hao công suất nhiều do quảng bá thông tin	-	+	Không cần thông tin quảng bá, hiệu quả giảm năng lượng tiêu hao
Khó triển khai cơ chế tự kiểm tra	-	+	Có thể tích hợp cơ chế tự kiểm tra

dụng từ các hệ thống khác miễn là chuẩn giao tiếp phù hợp với hệ thống hiện tại.

- **Kiến trúc truyền thông mạng trên chip** là thành phần chính thực hiện chức năng chuyển tiếp thông tin, được thiết lập bằng các bộ định tuyến và liên kết mạng giữa các bộ định tuyến.

- **Giao tiếp mạng** có chức năng kết nối các lõi IP vào kiến trúc truyền thông mạng trên chip để hoàn thiện hệ thống thành một thể thống nhất hoạt động ổn định và hiệu quả. Trong xu thế thiết kế lấy truyền thông làm trung tâm, các lõi IP được tái sử dụng từ các nguồn khác nhau thì có thể hoạt động theo nhiều chuẩn giao tiếp hiện có như: AMBA Bus, Avalon Bus, CoreConnect Bus, OCP, ... Do vậy, giao tiếp mạng cần được xây dựng chức năng thích ứng hoạt động truyền thông để kết nối chúng vào một môi trường truyền thông với cơ chế mạng thống nhất.

Kiến trúc truyền thông mạng trên chip là giải pháp truyền thông đáp ứng tối ưu cho các hệ thống trên chip đa chức năng phức hợp với nhiều ưu

điểm của giải pháp mạng so với các giải pháp truyền thông trên chip trước đây (Bảng 1.1). Truyền thông mạng trên chip trên cơ sở chia nhỏ đường truyền giữa các lõi IP theo nguyên tắc liên kết điểm-điểm và chia sẻ tài nguyên truyền thông một cách linh hoạt đã giải quyết được nhiều hạn chế của truyền thông dạng bus chia sẻ (đường truyền quá dài, băng thông hạn chế,...).

Như trong mô tả ở Hình 1.1, kiến trúc truyền thông mạng trên chip là tập hợp các bộ định tuyến được kết nối theo cấu trúc dạng lưới hai chiều 2D mesh tạo nên môi trường truyền thông cho phép các lõi IP trao đổi thông tin thông suốt và linh hoạt thông qua các giao tiếp mạng. Trong điều kiện trên chip, các bộ định tuyến được tối giản kiến trúc và chức năng để tiết giảm chi phí thiết kế và thực thi, do đó các giải pháp truyền thông cũng cần được tối ưu phù hợp với các ràng buộc của môi trường trên chip. Vậy nên, mạng trên chip đã và đang được nghiên cứu phát triển ngày càng đa dạng theo nhiều định hướng nhằm khắc phục các điểm hạn chế còn tồn tại cũng như gia tăng khả năng ứng dụng trong nhiều lĩnh vực.

Để đảm bảo khả năng truyền thông tin cậy và thông suốt, kiến trúc truyền thông mạng trên chip cũng được xây dựng với đặc điểm tổ chức của mạng truyền thống với hai phần chính như sau:

(A) ***Các thành phần vật lý*** là tập hợp các thành phần cơ bản thực hiện các hoạt động trao đổi thông tin bên trong môi trường mạng. Các thành phần vật lý được quản lý bởi tập hợp các giao thức truyền thông phù hợp để tạo thành môi trường mạng đáp ứng các yêu cầu truyền thông bên trong các hệ thống trên chip. Hình 1.1 mô tả mạng trên chip với các thành phần vật lý bao gồm: các bộ định tuyến, liên kết giữa các bộ định tuyến; và khối giao tiếp bộ định tuyến với các lõi IP.

- **Bộ định tuyến (*router*)** là thành phần chính trong mạng trên chip, thực hiện chức năng định tuyến chuyển tiếp thông tin từ nguồn đến đích thông qua các liên kết vật lý. Chức năng chính của bộ định tuyến là đệm thông tin, quản lý cấp phát và giải phóng tài nguyên đáp ứng cho các yêu cầu trao đổi thông tin từ các ứng dụng của hệ thống. Phụ thuộc vào yêu cầu truyền thông, cấu trúc mạng trên chip và các giao thức sử dụng, bộ định tuyến bảo đảm các yêu cầu kết nối từ bất kỳ ứng dụng nào trên mạng.
- **Liên kết vật lý (*physical link*)** là các đường dẫn tín hiệu lan truyền giữa các bộ định tuyến và giữa bộ định tuyến với các lõi IP. Đặc tính

của liên kết vật lý có sự phụ thuộc lớn vào công nghệ chế tạo, đặc điểm thiết kế và phương pháp thực thi của hệ thống [5].

Tập hợp các thành phần vật lý của mạng được bố trí sắp xếp theo những quy luật nhất định để đáp ứng tối ưu các yêu cầu của ứng dụng tạo nên cấu trúc liên kết (*topology*).

(B) ***Giao thức truyền thông*** là tập hợp các giao thức và giải thuật nhằm bảo đảm các kết nối giữa các lõi IP thông qua mạng trên chip và cân bằng các yêu cầu truyền thông nhằm đáp ứng tối ưu cho hoạt động của hệ thống. Giao thức truyền thông cần được xây dựng phù hợp với các thành phần vật lý đã được thiết kế và đặc điểm hoạt động của hệ thống trên chip.

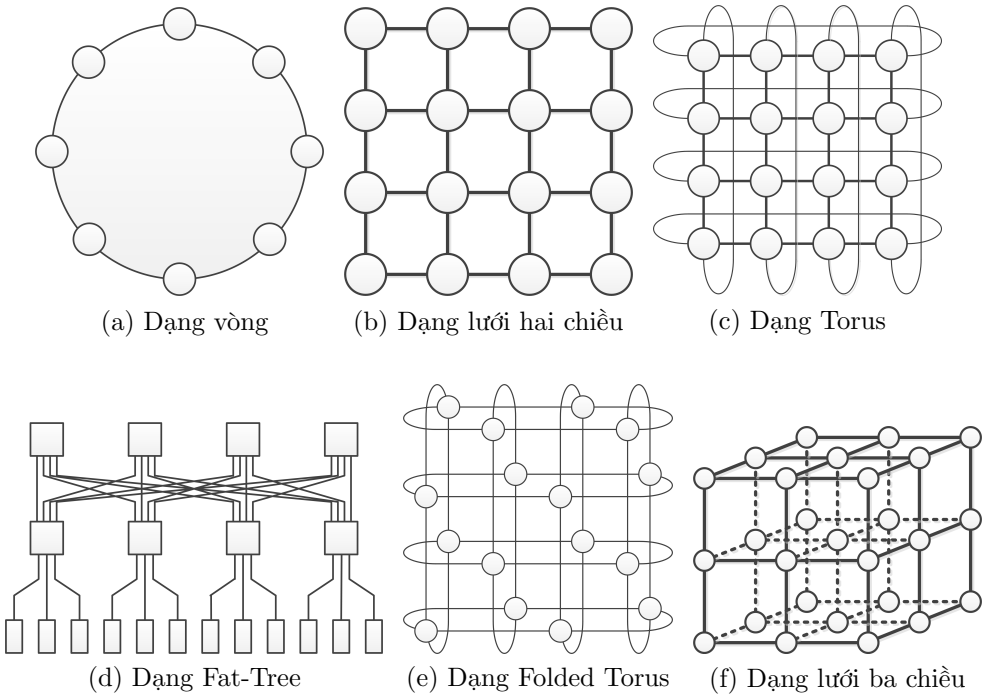
Theo cơ chế truyền thông của mạng, giao thức truyền thông tạo nên tập hợp các nguyên tắc trao đổi thông tin giữa các thành phần của hệ thống. Tập giao thức được phân lớp chức năng từ mức ứng dụng giao tiếp với người dùng hướng đến các quy định cụ thể cho lớp vật lý trong mô hình tham chiếu liên kết hệ thống mở (*OSI reference model*) [57].

1.2 Cấu trúc liên kết

Các thành phần bên trong của mạng truyền thông cần có sự tổ chức, sắp xếp theo một trật tự nhất định để tạo nên một hệ thống truyền thông phù hợp với yêu cầu của ứng dụng. Tập hợp các nguyên tắc để thiết lập các thành phần của mạng truyền thông được gọi là **cấu trúc liên kết (*topology*)**. Mạng trên chip là hướng áp dụng nguyên lý trao đổi thông tin qua môi trường mạng và có thể sử dụng hầu hết các dạng cấu trúc liên kết mạng truyền thống đã có như: dạng vòng, dạng lưới hai chiều (2D-mesh), dạng 2D-Torus, dạng hình cây (*Tree hay Fat Tree*), ... (Hình 1.2). Mỗi dạng cấu trúc sắp xếp đều có những ưu nhược điểm khác nhau, do vậy trong việc nghiên cứu và thiết kế cần có sự đáng giá tổng thể các cấu trúc liên kết khả dĩ để hướng đến một cấu trúc phù hợp cho mục tiêu mà hệ thống ứng dụng hướng đến.

Việc đánh giá hiệu quả hoạt động và khả năng tương thích của các cấu trúc liên kết với ứng dụng hệ thống và chi phí thiết kế thực hiện được thể hiện qua các tham số sau [29]:

1. **Cấp bậc bộ định tuyến (*router degree*)**: Là thông số chỉ số lượng liên kết của một bộ định tuyến với các bộ định tuyến khác trong mạng. Trong các cấu trúc liên kết, cấp bậc của mỗi bộ định tuyến có thể bằng nhau



(a) Dạng vòng (b) Dạng lưới hai chiều (c) Dạng Torus

(d) Dạng Fat-Tree (e) Dạng Folded Torus (f) Dạng lưới ba chiều

Hình 1.2: Một số dạng cấu trúc liên kết mạng trên chip.

hoặc khác nhau; trong mạng vòng (*ring*) cấp bậc định tuyến là bằng nhau, nhưng trong mạng lưới hai chiều 2D-mesh thì cấp bậc bộ định tuyến là khác nhau.

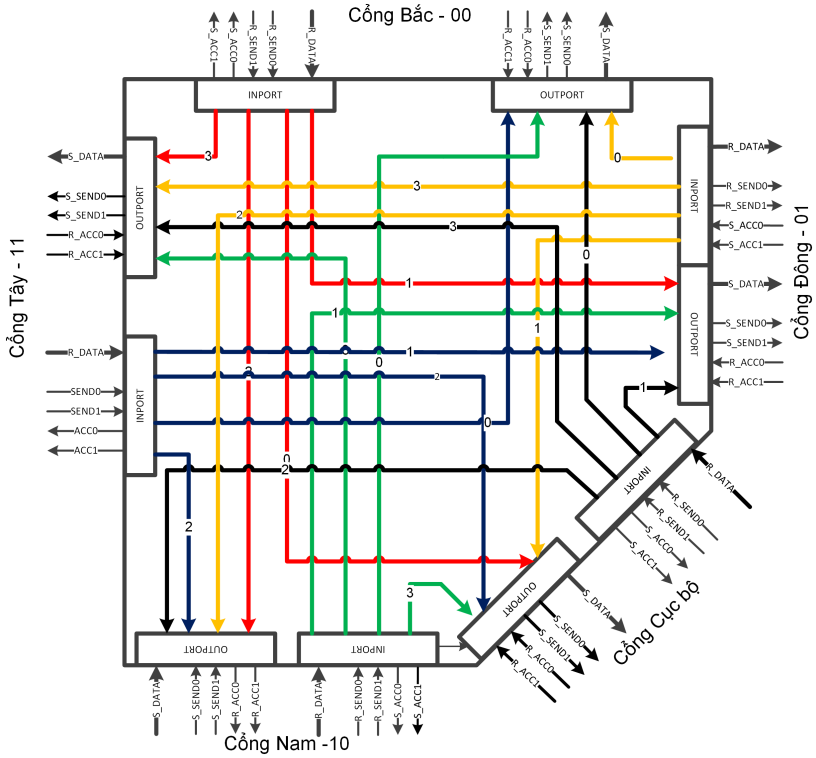
2. **Số lượng hop (*hop count*)**: Tham số này được tính cho từng đường định tuyến trong mạng, có giá trị bằng số lượng bộ định tuyến (*hop*) mà gói tin phải đi qua tính từ nguồn đến đích. Trên phương diện đánh giá hiệu năng toàn mạng, số lượng hop được tính trung bình cho tất cả các kết nối khả dĩ của bất kỳ cặp nguồn-đích nào trong mạng.

3. **Đường kính mạng (*network diameter*)**: Được xác định bằng đường định tuyến ngắn nhất giữa hai bộ định tuyến ở xa nhất trong mạng. Một cách tính khác thì đường kính mạng chính là giá trị cực đại của số lượng hop.

4. **Tính đối xứng (*symmetry*)**: Một cấu trúc mạng được xem là có tính đối xứng khi trong mạng có tâm đối xứng (*automorphism*) cho phép ánh xạ một nút mạng A thành một nút mạng B qua điểm đối xứng [14].

Về nguyên tắc, mạng trên chip được nghiên cứu với nhiều dạng cấu trúc liên kết và có thể sử dụng tất cả các cấu trúc liên kết của mạng máy tính. Ngoài ra, một số ứng dụng cụ thể còn sử dụng các mạng bất quy tắc (*irregular topology*) để nâng cao hiệu quả thiết kế và bảo đảm tốt hơn các yêu cầu truyền

thông của hệ thống trên chip. Tuy nhiên, phần lớn các mạng trên chip sử dụng cấu trúc liên kết dạng lưới hai chiều để phù hợp hơn với công nghệ chế tạo vi mạch hiện tại [13]. Mô hình mạng sử dụng cấu trúc liên kết dạng lưới hai chiều (Hình 1.2b) sử dụng các bộ định tuyến đồng dạng với kiến trúc gồm 5 cổng vào/ra theo các hướng: Bắc (*North*), Đông (*East*), Nam (*South*), Tây (*West*) và một cổng kết nối với lõi IP được gọi là Cổng cục bộ (*Local port*) (như được mô tả trong Hình 1.3). Một biến thể khác của dạng mạng lưới hai chiều 2D-mesh là cấu trúc liên kết *torus* và *folded torus* (Hình 1.2c và 1.2e). Trong cấu trúc liên kết *torus*, các bộ định tuyến có cùng bậc và hỗ trợ linh hoạt các đường kết nối trong mạng; tuy nhiên giải thuật định tuyến cần thích ứng với cấu trúc này sẽ phức tạp hơn so với giải thuật định tuyến trong dạng lưới hai chiều 2D-mesh.



Hình 1.3: Cấu trúc của một bộ định tuyến dùng trong mạng 2D-mesh [34].

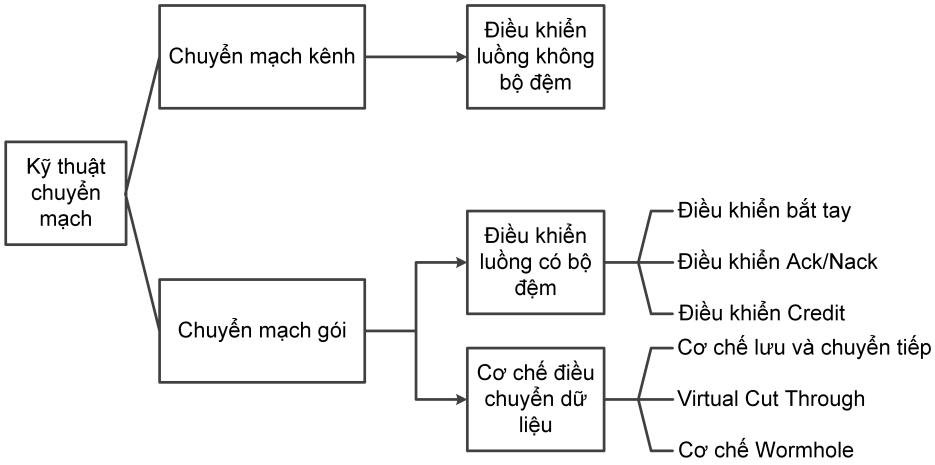
Ngoài ra, một số nghiên cứu về mạng trên chip sử dụng các dạng cấu trúc liên kết khác như Guerrier và cộng sự đã sử dụng cấu trúc liên kết dạng Fat-Tree [25], hoặc sử dụng các cấu trúc liên kết bất quy tắc để thích ứng với các ứng dụng trong những hệ thống cụ thể [45, 61]. Khi sử dụng cấu trúc liên kết bất quy tắc, chúng ta cần xây dựng các giao thức truyền thông chuyên biệt để thích ứng với từng cấu trúc cụ thể của mạng. Điều này sẽ tối ưu cho ứng dụng của hệ thống được thiết kế, nhưng sẽ khó sử dụng lại thiết kế cho các ứng dụng hoặc hệ thống khác.

Cấu trúc liên kết mạng là cơ sở quan trọng cho mọi nghiên cứu tìm hiểu

hoạt động truyền thông mạng và mạng trên chip nói riêng. Với đa dạng các cấu trúc mạng có thể áp dụng cho mạng trên chip; tuy nhiên, cấu trúc liên kết dạng lưới được tập trung nghiên cứu cho mạng trên chip với nhiều đặc điểm thuận lợi. Do vậy, luận án này tập trung chủ yếu cho việc nghiên cứu hoạt động truyền thông mạng trên chip với cấu trúc liên kết dạng lưới hai chiều 2D-mesh [15].

1.3 Kỹ thuật truyền thông

Về cơ bản mạng trên chip vẫn tồn tại hai xu thế truyền thông sử dụng kỹ thuật chuyển mạch kênh (*circuit switching*) và kỹ thuật chuyển mạch gói (*packet switching*). Nhưng với đặc điểm chia sẻ tài nguyên linh hoạt và khả năng đáp ứng đa dạng các nhu cầu truyền thông, mạng trên chip sử dụng chuyển mạch gói được quan tâm nghiên cứu và phát triển mạnh trong thời gian qua. Hoạt động chuyển mạch gói của mạng trên chip cần kết hợp với các kỹ thuật truyền thông khác để đảm bảo hoạt động trao đổi thông tin là tin cậy và hiệu quả như: cơ chế điều khiển luồng (*flow control mechanism*); cơ chế điều chuyển dữ liệu (*data switching flow control*) và nguyên tắc tổ chức bộ đệm bên trong bộ định tuyến.

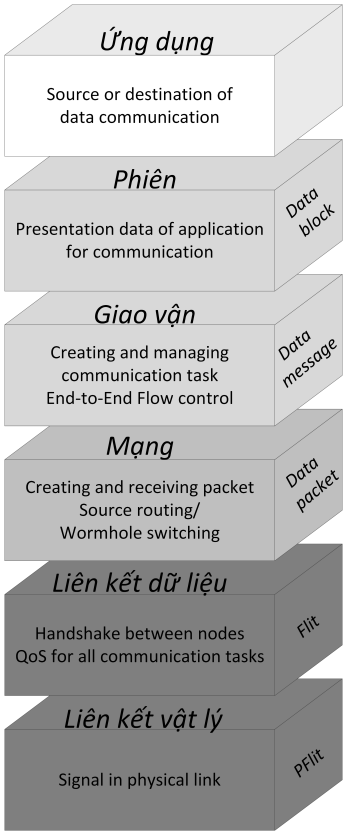


Hình 1.4: Kỹ thuật truyền thông cho mạng trên chip.

Hoạt động trao đổi đảm bảo các giao dịch truyền dữ liệu tin cậy trong hệ thống cần có các giao thức truyền thông phù hợp với hạ tầng truyền thông và đáp ứng yêu cầu thông tin của toàn hệ thống. Giao thức truyền thông bảo đảm các thành phần trong hệ thống trao đổi dữ liệu một cách tin cậy và hiệu quả giữa các khối chức năng khác nhau thông qua hạ tầng mạng. Với xuất phát điểm từ mô hình mạng máy tính, mạng trên chip vẫn sử dụng mô hình tham chiếu liên kết hệ thống mở với một số thay đổi cho phù hợp điều kiện thực thi trên vi

mạch như trong Hình 1.5 [57]. Trên cơ sở kỹ thuật chuyển mạch gói của mạng trên chip, thông tin được tạo lập từ ứng dụng thường có dạng khối lớn (hoặc luồng dữ liệu kéo dài theo thời gian thực) sẽ được sắp xếp và đóng gói phù hợp với các giao thức truyền thông được sử dụng ở các lớp tương ứng trong mô hình phân lớp. Khối dữ liệu (*data block*) được tổ chức hướng đến hoạt động trao đổi thông tin cần được chia nhỏ thành các bản tin (*message*) tương ứng trong hoạt động truyền thông được thiết lập giữa các lõi IP bên trong hệ thống. Với chức năng thích ứng giao tiếp của mọi khối lõi IP, khối giao tiếp mạng cần đóng gói các dữ liệu đa dạng ở trên thành gói tin (*data packet*) có định dạng phù hợp với khả năng truyền thông của mạng trên chip; trong đó có sự kết hợp của dữ liệu cần trao đổi với thông tin điều khiển dùng cho mục đích truyền thông.

Mạng trên chip thực hiện chức năng truyền thông bằng khả năng chuyển tải các gói tin đã được đóng gói tại giao tiếp mạng và phân nhỏ thành các đơn vị điều khiển luồng (flit: *flow control unit*). Đơn vị thông tin flit là độ rộng của từ được sử dụng trong các hoạt động truyền thông giữa các thành phần của mạng trên chip, hiện nay mạng trên chip thường sử dụng đơn vị *flit* có độ rộng dữ liệu 32 *bit* hoặc 64 *bit*.



Hình 1.5: Mô hình phân lớp áp dụng cho mạng trên chip.

Dựa vào đặc điểm của mạng trên chip và sự phân cấp theo các lớp chức

năng trình bày trình Hình 1.5, môi trường truyền thông của mạng trên chip tập trung ở ba lớp hướng mạng gồm: lớp liên kết vật lý (*Physical link*), lớp liên kết dữ liệu (*Data link*) và lớp mạng (*Network*). Các lớp mức cao hướng đến hoạt động của các ứng dụng chỉ sử dụng các dịch vụ được cung cấp từ các lớp dưới trong quá trình truyền thông.

1. **Lớp liên kết vật lý**

   Lớp liên kết vật lý mô tả các tiêu chuẩn giao tiếp giữa hai hoặc nhiều hơn các thành phần vật lý của mạng tạo nên nền tảng hạ tầng mạng, là cơ sở tạo nên tài nguyên truyền thông của mạng. Liên kết vật lý cung cấp khả năng trao đổi thông tin ở dạng tín hiệu lan truyền trong môi trường vật lý giữa các thành phần mạng; nên còn được gọi là phương tiện truyền thông. Ở mạng trên chip, lớp liên kết vật lý thể hiện nguyên lý đồng bộ và định dạng tín hiệu lan truyền bên trong mạng, giữa các bộ định tuyến, giữa bộ định tuyến và giao tiếp mạng hoặc khối chức năng ứng dụng. Thông tin truyền ở lớp này được mô tả là tập các tín hiệu đơn vị (PFlit: *Physical flit*) là nhóm tín hiệu logic tương ứng với từ xử lý của hệ thống mạng.

2. **Lớp liên kết dữ liệu**

   Lớp liên kết dữ liệu thực hiện hoạt động trao đổi thông tin đơn vị giữa các thành phần bên trong mạng trên nền tảng liên kết vật lý đã có giữa các bộ định tuyến (Hình 1.5). Lớp liên kết dữ liệu quản lý tài nguyên truyền thông để bảo đảm các giao tác là tin cậy và thông suốt trong toàn mạng. Các giao thức lớp liên kết dữ liệu quy định hoạt động bắt tay trao đổi thông tin giữa các thành phần mạng.

   Mạng trên chip với đặc điểm tinh giảm, các bộ định tuyến thực hiện chức năng liên kết điểm-điểm giữa các bộ định tuyến và tạo nên hạ tầng kết nối cho các giao tác truyền thông giữa các ứng dụng ở các lớp mức cao hơn. Phụ thuộc vào định hướng thiết kế và yêu cầu của hệ thống, bộ định tuyến có thể kết hợp với thông tin định tuyến ở lớp mạng để thực hiện quá trình thiết lập đường đi của thông tin trong các giao tác truyền thông. Hoạt động trao đổi dữ liệu ở mạng trên chip dựa trên đơn vị thông tin cơ bản là flit.

3. **Lớp mạng**

   Lớp mạng tạo nên các dịch vụ và chức năng truyền thông cho phép cung cấp khả năng trao đổi dữ liệu giữa các ứng dụng thông qua cơ sở hạ tầng mạng đã có ở hai lớp dưới.

Trong mạng trên chip, lớp mạng thực hiện chức năng đóng/mở gói tin thích ứng với tổ chức hạ tầng truyền thông ở các lớp hướng mạng. Đồng thời trong một số các trường hợp cụ thể thực hiện chức năng định tuyến đầu-cuối (*end-to-end*).

4. **Lớp giao vận**
   Lớp giao vận cung cấp dịch vụ giao vận thông suốt cho các lớp hướng ứng dụng thông qua các giao thức điều khiển luồng tin vào mạng cũng như bảo đảm độ tin cậy của các giao tác truyền thông. Chức năng chính của lớp giao vận là thiết lập, duy trì hoạt động truyền dẫn và giải phóng các kết nối thông suốt qua mạng từ nguồn tin hướng đến đích. Do đó, lớp giao vận cho phép cách ly các lớp hướng mạng và các lớp hướng ứng dụng; điều này cho phép phát triển các ứng dụng đa dạng mà không cần quan tâm đến hoạt động truyền thông cụ thể.

5. **Lớp phiên**
   Lớp phiên biểu diễn thông tin từ lớp ứng dụng theo các nguyên tắc phù hợp với hoạt động truyền thông của hệ thống, làm cơ sở thực hiện các phiên làm việc trong hoạt động trao đổi thông tin giữa các ứng dụng. Đồng thời, trong hoạt động truyền thông lớp phiên có nhiệm vụ cung cấp khả năng quản lý các phiên làm việc và chế độ trao đổi thông tin tin cậy giữa hai đối tượng truyền thông (*song công, bán song công*).

6. **Lớp ứng dụng**
   Lớp ứng dụng là lớp cao nhất của mô hình phân lớp với nhiệm vụ giao tiếp người dùng, cung cấp khả năng ứng dụng của hệ thống và cũng là đích đến của thông tin trao đổi qua mạng.

1.3.1 Cơ chế điều khiển luồng

Cơ chế điều khiển luồng là cơ sở cho nguyên tắc phân bổ tài nguyên truyền thông: băng thông (*bandwidth*), dung lượng bộ đệm (*buffer capacity*), trạng thái điều khiển, ...; để bảo đảm các giao tác truyền thông qua mạng là tin cậy [1]. Điều khiển luồng dựa vào đặc tính cấu trúc và kỹ thuật chuyển mạch được sử dụng để phân làm hai loại cơ chế chính là điều khiển luồng không bộ đệm (*bufferless*) và điều khiển luồng có bộ đệm (*buffer*).

- **Điều khiển luồng không bộ đệm**
  Là cơ chế điều khiển luồng kết hợp với kỹ thuật chuyển mạch kênh với các

thành phần nút mạng trung gian không có bộ đệm thông tin. Hoạt động truyền thông dựa trên nguyên tắc bắt tay tạo kết nối trước mới tiến hành truyền tin; do đó thông tin truyền từ nguồn đến đích trên các liên kết dành riêng mà không cần bộ đệm ở các nút trung gian. Tuy nhiên, cơ chế bắt tay kết nối trước phải duy trì tài nguyên được cấp phát trong toàn phiên làm việc. Do đó, cơ chế này có hạn chế lớn là hiệu quả sử dụng tài nguyên truyền thông thấp.

- **Điều khiển luồng có bộ đệm**
  Trong mạng truyền thông sử dụng kỹ thuật chuyển mạch gói, thông tin sẽ đi qua các nút mạng trung gian cần được lưu trữ trước khi chuyển đến nút mạng tiếp theo. Do đó, cần có cơ chế điều khiển luồng sử dụng việc điều khiển bộ đệm để thực hiện quá trình chuyển thông tin qua mạng. Điều khiển luồng có bộ đệm tạo nên sự linh hoạt trong hoạt động cấp phát tài nguyên, cân bằng yêu cầu truyền thông.

Với hướng mạng trên chip, các nghiên cứu hiện nay sử dụng rất nhiều kỹ thuật, cơ chế khác nhau do vậy cơ chế điều khiển luồng là rất đa dạng. Hướng đến mô hình mạng sử dụng chuyển mạch gói cho phép linh hoạt trong truyền thông, nên các cơ chế điều khiển luồng có bộ đệm đang rất được quan tâm nghiên cứu và triển khai như: điều khiển Ack/Nack, điều khiển luồng Handshaking.

Điều khiển luồng thông qua tín hiệu bắt tay (*handshaking signals*): Cơ chế điều khiển này dựa trên việc xác thực thông qua các tín hiệu bắt tay giữa khối phát và khối thu. Ở đầu phát, tín hiệu xác thực (*valid*) được phát kèm thông tin và đầu nhận sẽ chèn trả lời xác nhận vào tín hiệu xác thực và xóa thông tin vừa nhận. Một số mạng trên chip điển hình sử dụng cơ chế điều khiển hướng bắt tay như trong công trình [70].

Điều khiển luồng Ack/Nack: Cơ chế này bắt buộc phải sao lưu thông tin phát và chỉ được xóa khi nhận được tín hiệu xác nhận (*Ack*) từ bên nhận. Trong trường hợp bên phát không nhận được xác nhận hoặc nhận được tín hiệu không xác nhận (*Nack*); bên nhận sẽ tự động phát lại thông tin đã phát. Cơ chế này được Bertozzi và cộng sự sử dụng trong mạng trên chip với kiến trúc xPIPES [4].

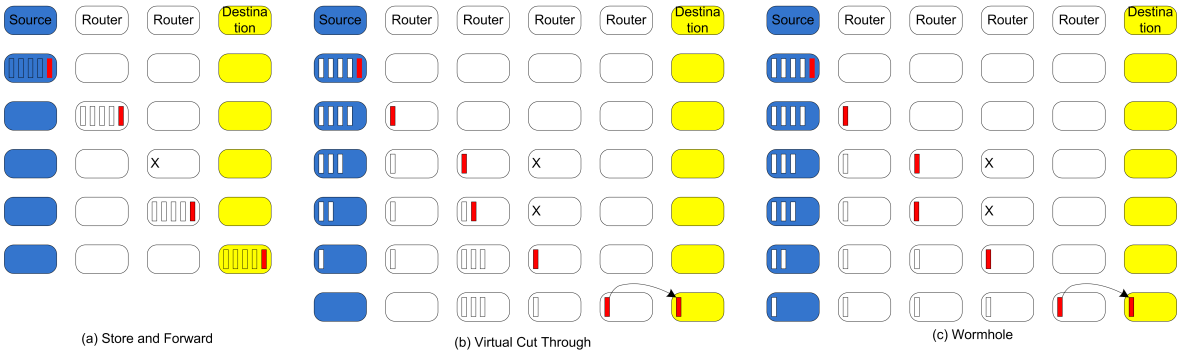
Điều khiển luồng dựa vào *credit*. Kết hợp ưu điểm của hai cơ chế trên, điều khiển luồng dựa vào ***credit*** thực hiện điều khiển quá trình trao đổi thông tin khi phía nhận còn khả năng đáp ứng, được gọi là ***Credit***. Nguyên tắc điều khiển luồng dựa vào credit được tóm lược như sau:

- Phía nhận: giá trị ***credit*** sẽ giảm đi 1 khi tiếp nhận một đơn vị thông tin. Hoạt động nhận sẽ tạm dừng khi giá trị credit giảm về 0.
- Phía phát: chỉ phát tin khi đầu nhận có giá trị credit khác 0. Ở các nút mạng sử dụng kỹ thuật điều khiển luồng dựa vào ***credit*** sau mỗi lần gói tin được phát đi thì giá trị ***credit*** tại đầu nhận lại tăng lên 1 giá trị và sẵn sàng nhận tiếp gói tin. Cơ chế điều khiển luồng credit-based hiện nay đang được sử dụng trong nhiều nghiên cứu về mạng trên chip như trong các kiến trúc ANoC [2] và QNoC [8].

Ngoài ra, hoạt động điều khiển luồng của mạng trên chip vẫn có thể sử dụng các cơ chế của mạng máy tính là cơ chế **Dừng và đợi (*Stop and Wait*)** và cơ chế **Cửa sổ trượt (*Sliding Window*)**. Tuy nhiên, các cơ chế này đòi hỏi sử dụng bộ đệm lớn. Đây là một hạn chế khi thực hiện trên hệ thống trên chip do chi phí thiết kế và thực thi.

1.3.2 Cơ chế điều chuyển dữ liệu

Cơ chế điều chuyển dữ liệu là nguyên tắc chuyển giao thông tin trong các giao tác truyền thông giữa các nút mạng trên đường đi của thông tin từ nguồn đến đích. Cơ chế điều chuyển dữ liệu được đề xuất để đảm bảo chất lượng dịch vụ (QoS: *Quality of Service*) cho các giao tác truyền thông trong điều kiện ràng buộc của tài nguyên truyền thông giới hạn của môi trường mạng. Các cơ chế điều chuyển dữ liệu được áp dụng trong kỹ thuật chuyển mạch gói với ba cơ chế thường gặp: cơ chế lưu và chuyển tiếp (SAF: *Store and Forward*), cơ chế xuyên cắt ảo (VCT: *Virtual Cut-Through*) và cơ chế WH (*Wormhole*) (Hình 1.6).



Hình 1.6: Các cơ chế điều chuyển dữ liệu.

Cơ chế lưu và chuyển tiếp (Hình 1.6a). Hoạt động trên cơ sở chuyển mạch gói, nút mạng nhận và lưu đầy đủ gói tin vào bộ đệm trước khi chuyển tiếp tới nút mạng tiếp theo. Hoạt động lưu và chuyển tiếp gói tin sẽ luôn bảo

đảm độ tin cậy giao tác truyền tin. Tuy nhiên, cơ chế này cần tài nguyên bộ đệm lớn hơn kích thước gói tin cực đại trong tất cả các bộ định tuyến và thời gian chờ chuyển tiếp giữa các bộ định tuyến trung gian sẽ làm gian tăng độ trễ truyền của mạng.

Cơ chế VCT (Hình 1.6b). Nhược điểm của cơ chế SAF là thời gian chuyển tiếp gói tin ở các nút mạng trung gian bị kéo dài do phải nhận đủ gói tin thì các nút mạng trung gian mới chuyển tiếp đến nút mạng tiếp theo. Để khắc phục nhược điểm này, cơ chế VCT cho phép các nút mạng chuyển tiếp ngay thông tin sau khi nhận được nếu các nút mạng liền sau trên đường định tuyến sẵn sàng nhận tin. Tuy nhiên, trong quá trình chuyển tiếp gói tin đến đích, nếu có một nút trung gian không thể nhận tin thì gói tin sẽ được lưu lại tại nút mạng ngay liền trước. Do vậy, mặc dù giảm được độ trễ truyền của mạng, cơ chế này vẫn yêu cầu các nút mạng có đủ bộ đệm cho những gói tin có kích thước lớn nhất theo giao thức truyền thông quy định.

Cơ chế Wormhole (Hình 1.6c). Kết hợp các ưu điểm và khắc phục nhược điểm của hai cơ chế SAF và VCT, cơ chế WH cho phép chuyển tiếp thông tin ngay khi nút mạng liền sau có thể nhận thông tin nhưng lại không tiếp nhận thông tin mới nếu chưa chuyển tiếp được thông tin đang lưu giữ tại nút mạng. Do đó, thông tin trao đổi của mỗi giao tác sẽ trải đều trên đường đi của gói tin, và trong trường hợp đơn vị tin đầu tiên bị chặn thì toàn bộ gói tin sẽ chiếm dụng tài nguyên của các nút mạng trung gian như trong (Hình 1.6c). Cơ chế này sẽ giảm được cả thời gian trễ truyền qua các nút mạng trung gian và độ sâu của bộ đệm tại các nút mạng. Nhược điểm của cơ chế này là có thể gây nên hiện tượng khóa vòng chết (*deadlock*) do các giao tác truyền thông chiếm dụng tài nguyên lẫn nhau mà không thể giải phóng được. Hiện tượng khóa vòng này sẽ được trình bày cụ thể tại Mục 1.4.

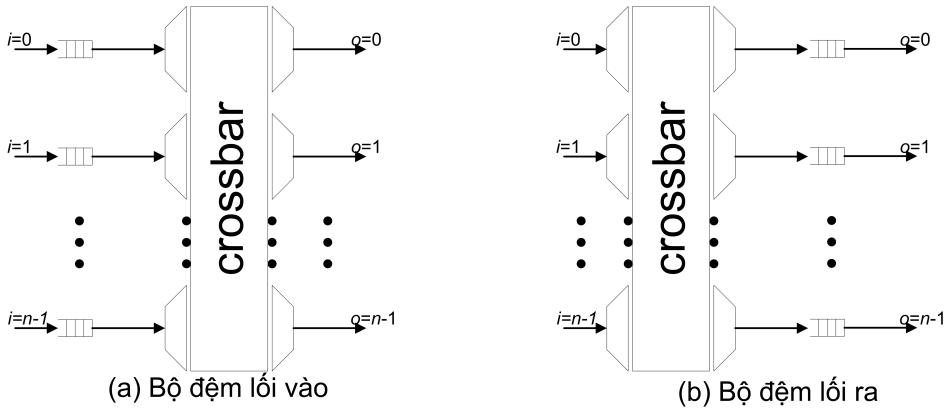
1.3.3 Chiến lược bộ đệm

Hoạt động chuyển tin trong mạng chuyển mạch gói có được sự linh hoạt nhờ vào khả năng chia sẻ tài nguyên của các giao tác truyền thông một cách tin cậy và hiệu quả dựa trên nhiều kỹ thuật kết hợp. Hoạt động chuyển tiếp gói tin trong mạng không chiếm giữ tài nguyên liên tục trong suốt quá trình trao đổi giữa các đối tượng mà chỉ chuyển tin dưới dạng gói thông qua các điểm trung gian hướng đến đích. Tuy nhiên, gói tin đi qua các điểm trung gian cần được xác thực và xử lý chuyển tiếp để bảo đảm độ tin cậy của các giao tác truyền thông. Do đó, mỗi bộ định tuyến trong mạng cần có bộ đệm với kích thước phù hợp để

lưu giữ thông tin một cách tối ưu và tin cậy. Chiến lược bộ đệm là tập hợp các nguyên lý để xác định giải pháp lưu trữ tại bộ định tuyến gồm: vị trí bộ đệm và kích thước bộ đệm trong mỗi bộ định tuyến. Dù hiện nay các nghiên cứu mạng trên chip đã đề xuất rất nhiều kiến trúc bộ định tuyến khác nhau, nhưng về cơ bản luôn có một số thành phần chính: khối nhớ đệm, ma trận tín hiệu nối chéo (*crossbar*) và một số khối chức năng điều khiển. Dựa vào đặc điểm kiến trúc bộ định tuyến khi sắp xếp vị trí bộ đệm sẽ tạo nên các chiến lược bộ đệm phù hợp. Trong lĩnh vực nghiên cứu mạng trên chip hiện nay tồn tại 4 chiến lược bộ đệm chính [67] gồm: hàng đợi lối vào (*input queuing*), hàng đợi lối ra (*output queuing*), hàng đợi lối ra ảo (*virtual output queuing*) và hàng đợi lối vào ưu tiên kênh ảo (*virtual channel priority input queuing*).

Để thuận lợi cho việc mô tả đặc điểm chung của tất cả các chiến lược bộ đệm trong phần sau đều áp dụng chung cho bộ định tuyến với kiến trúc n cổng lối vào, n cổng lối ra và một ma trận tín hiệu nối chéo $n \times n$:

- **Hàng đợi lối vào**: Giải pháp bộ đệm cơ bản nhất là tạo ra n bộ đệm tương ứng với n lối vào trong mỗi bộ định tuyến (Hình 1.7a). Với bộ định tuyến dạng này sẽ lưu ngay thông tin đến ở mỗi đầu vào và chờ sự phân xử để chuyển hướng đến lối ra phù hợp. Chiến lược này bảo đảm thông tin không bị mất khi có tranh chấp đầu ra. Tuy nhiên, hiệu quả sử dụng truyền thông không cao do hiệu ứng khóa đầu luồng (HOL: *head-of-line blocking*).

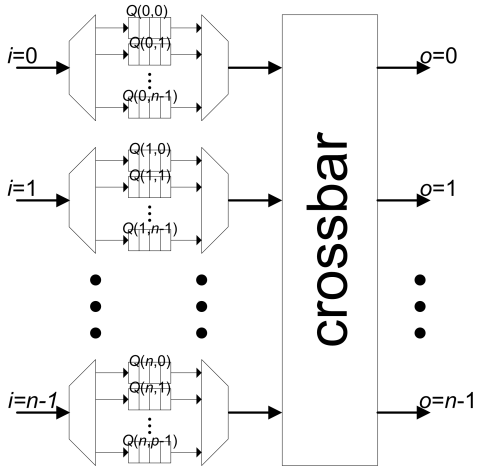


Hình 1.7: Chiến lược bộ đệm lối vào và lối ra.

Hiệu ứng HOL xảy ra khi một giao tác chiếm vị trí đầu hàng đợi không thể chuyển đến hướng ra, lúc đó tất cả các thông tin trao đổi khác đều cùng bị chặn lại. Hiệu ứng HOL có ảnh hưởng lớn đến hiệu quả truyền thông, và càng trở nên trầm trọng khi cấp bậc bộ định tuyến (n) càng lớn.

- **Hàng đợi lối ra**: các bộ định tuyến sử dụng chiến lược hàng đợi lối ra

sẽ bố trí các bộ đệm ở lối ra, sau khi thông tin chuyển qua ma trận tín hiệu nối chéo (*crossbar*) (Hình 1.7b). Với vị trí này, tất các thông tin nhận được tại các khối cổng lối vào sẽ được định tuyến ngay và chuyển hướng phù hợp mới được lưu trữ tại khối cổng lối ra. Do thông tin từ cổng lối vào được chuyển ngay đến bộ đệm tại lối ra nên hạn đến được hiện tượng HOL. Tuy nhiên, chiến lược này lại yêu cầu khối ma trận tín hiệu nối chéo có tốc độ kết nối đáp ứng cao (gấp n lần so với tốc độ chuyển tin của bộ định tuyến). Trong điều kiện thực thi trên chip thì việc yêu cầu một khối chức năng hoạt động ở tốc độ cao so với các khối khác là một hạn chế rất lớn.



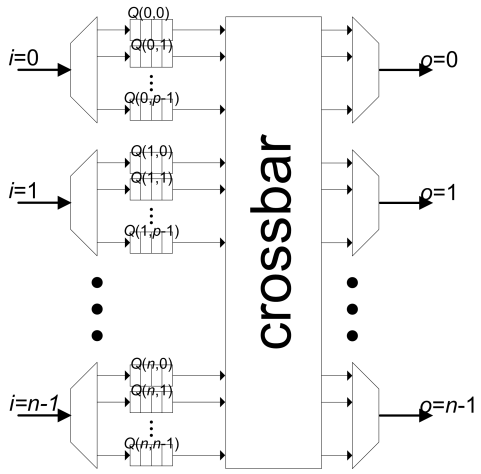
Hình 1.8: Chiến lược bộ đệm lối ra ảo.

- **Hàng đợi lối ra ảo (VOQ)**: Chiến lược này đặt bộ đệm đặt ở sau khối cổng lối vào với số lượng là n^2 cho n cổng lối vào của bộ định tuyến như trong Hình 1.8. Với kiến trúc bộ đệm được phân xử ngay tại lối vào nhưng tăng kích thước bộ đệm cho phép kết hợp cả hai ưu điểm của hai chiến lược bộ đệm vừa nêu. Chiến lược VOQ vừa có khả năng loại bỏ hiệu ứng HOL và cũng không yêu cầu khối ma trận nối chéo hoạt động ở tốc độ cao. Do đó chiến lược VOQ cho phép thông tin lối vào được xử lý và chuyển hướng đến lối ra cùng thời điểm, điều này làm tăng hiệu quả truyền thông.

  Tuy nhiên, hạn chế lớn nhất của chiến lược này là chi phí thiết kế thực thi cho kiến trúc bộ đệm trong mỗi bộ định tuyến tăng nhanh theo bậc bộ định tuyến. Để đáp ứng được hoạt động đệm dữ liệu, chúng ta cần n^2 bộ đệm cho mỗi bộ định tuyến trong mạng.

- **Hàng đợi lối vào ưu tiên kênh ảo (VCPIQ)** là một chiến lược bộ đệm cải tiến của chiến lược VOQ nhằm tăng hiệu năng truyền thông nhưng hạn

chế chi phí thực thi của VOQ. Kiến trúc bộ định tuyến sử dụng chiến lược bộ đệm VCPIQ được mô tả cơ bản trong Hình 1.9.



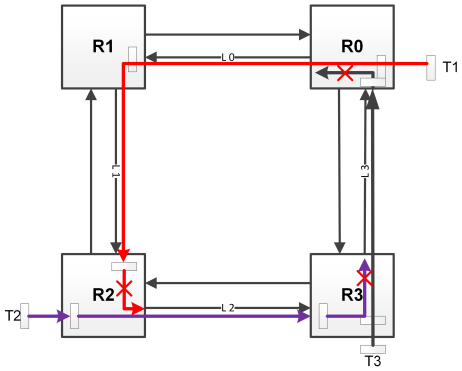
Hình 1.9: Chiến lược bộ đệm lối vào ưu tiên kênh ảo.

Trong chiến lược bộ đệm VCPIQ, mỗi cổng lối vào cho một kênh vật lý sử dụng p bộ đệm ($p < n$) tương ứng với p kênh ảo. Dựa trên nguyên lý hoạt động của chiến lược VOQ nên giải pháp này hạn chế được hiện tượng HOL và giảm được số lượng bộ đệm để giảm chi phí trong thực thi trên chip. Trong nghiên cứu đánh giá các chiến lược bộ đệm, Pande và cộng sự đã chỉ ra được số lượng kênh ảo tối ưu bằng 4 ($p = 4$) để đạt được cân bằng giữa chi phí và hiệu quả truyền thông [53].

1.4 Giải thuật định tuyến

Hoạt động truyền thông đặt ra yêu cầu tiên quyết là thông tin phải được định tuyến từ nguồn đến đích một cách chính xác và tin cậy. Với đặc điểm đa dạng của môi trường mạng, thông tin từ nguồn đến đích có thể được chuyển qua nhiều điểm trung gian linh hoạt tạo nên các đường định tuyến (*routing path*). Đáp ứng yêu cầu kết nối, mạng cần có tập các nguyên tắc để quyết định đường đi của thông tin từ nguồn đến đích tin cậy và hiệu quả dựa trên khả năng đáp ứng của hạ tầng mạng đang có. Tập các nguyên tắc xác định đường đi trong mạng được gọi là giải thuật định tuyến (*routing algorithm*). Giải thuật định tuyến cần bảo đảm kết nối tin cậy cho bất kỳ cặp nguồn-đích trong mạng và cân bằng lưu lượng thông tin dọc theo các liên kết khả dĩ đang có của môi trường mạng. Đồng thời, giải thuật định tuyến cần được xây dựng phù hợp với đặc điểm của hệ thống, hiệu quả cho tất cả các giao tác truyền thông của hệ thống đáp ứng cao nhất các yêu cầu chất lượng dịch vụ.

Hiện tượng khóa vòng chết (deadlock) là hiện tượng các gói tin của các giao tác truyền thông, trong mạng liên tục truy vấn tài nguyên truyền thông tạo thành một vòng kín không thể phân xử và giải phóng được tài nguyên truyền thông. Khi có hiện tượng khóa vòng, đường định tuyến của các giao tác tạo nên vòng kín và truy vấn tài nguyên lẫn nhau mà không thể truyền đến đích của mỗi giao tác. Hình 1.10 mô tả một trường hợp khóa vòng giữa ba giao tác truyền thông (T1, T2 và T3).



Hình 1.10: Hiện tượng khóa vòng.

Trong Hình 1.10, hiện tượng khóa vòng chết xảy ra do ba giao tác truyền thông T1, T2 và T3 chiếm dụng tài nguyên truyền thông lẫn nhau và không thể giải phóng được. Quá trình tranh chấp tài nguyên của giao tác T1 và T2 xảy ra tại bộ định tuyến R2 và tranh chấp của giao tác T3 và T1 xảy ra tại bộ định tuyến R0 tạo nên vòng khóa kín không thể phân xử. Do đó, cả ba giao tác T1, T2 và T3 không thể hoàn thành và giải phóng tài nguyên, sẽ ảnh hưởng lớn đến hoạt động toàn hệ thống.

Hiện tượng khóa vòng sống (*livelock*) là hiện tượng các gói tin không thể tiến đến đích của giao tác ngay cả khi gói tin không bị chặn. Trong trường hợp gói tin không hướng được đến đích sẽ liên tục chuyển hướng trong mạng và tạo nên hoạt động nhân tải mạng, tác động đến nhiều giao dịch khác cũng như khả năng đáp ứng truyền thông của mạng.

1.4.1 Phân loại định tuyến

Giải thuật định tuyến đã và đang được nghiên cứu nhiều trong mạng truyền thống cũng như riêng cho mạng trên chip. Nguyên lý định tuyến và các tiêu chí đánh giá giải thuật định tuyến cũng rất đa dạng [1, 29]. Việc sắp xếp và đánh giá các giải thuật định tuyến có thể dựa trên các tiêu chí như sau:

1. **Khả năng thích nghi trong chọn đường định tuyến**

Trong môi trường mạng, hoạt động giao tác truyền thông có thể đi theo nhiều đường định tuyến hướng từ nguồn đến đích. Các giải thuật định tuyến tạo nên cơ sở quyết định đường đi của thông tin có thể cố định để bảo đảm cân bằng tải mạng hoặc linh hoạt để đáp ứng các điều kiện hoạt động của mạng. Dựa vào đặc tính này có thể phân các giải thuật định tuyến làm hai loại chính: là định tuyến xác định (*deterministic routing*) và định tuyến thích nghi (*adaptive routing*).

2. **Nguyên lý quyết định định tuyến**.
   Giải thuật định tuyến sử dụng giải pháp tạo nên các thông tin định tuyến để quyết định đường đi của giao tác bởi nhiều cách khác nhau. Trong môi trường mạng trên chip, việc quyết định đường định tuyến có thể bằng máy trạng thái hữu hạn (FSM: *Finite State Machine*) hoặc sử dụng bảng tìm kiếm (LUT: *Lookup-Table*). Bên cạnh đó, một số công trình sử dụng các mạch logic tổ hợp làm cơ sở để xác định định tuyến [18].

3. **Thực thi định tuyến**
   Hoạt động định tuyến trong mạng được thực hiện trên cơ sở giải thuật định tuyến kết hợp với nguyên lý quyết định để tạo nên thông tin định tuyến. Việc thực thi định tuyến phụ thuộc vào đặc điểm của hệ thống và cân bằng với các điều kiện hoạt động của hệ thống. Phân loại dựa trên đặc điểm thực hiện quá trình định tuyến có thể chia các giải thuật định tuyến thành: định tuyến tại nguồn (*source routing*), định tuyến phân tán (*distributed routing*) và định tuyến trung tâm (*centrallized routing*). Ngoài ra, một số nghiên cứu mạng trên chip sử dụng kết hợp giải pháp thực thi định tuyến tạo nên dạng định tuyến nhiều đoạn (*multiphase routing*).

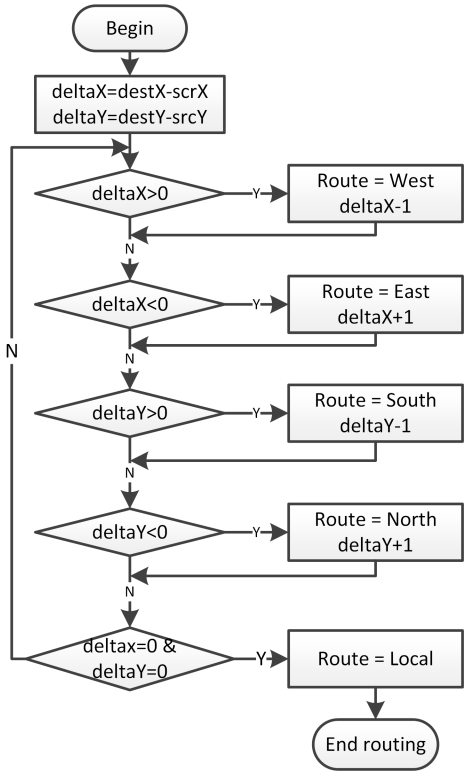
Với khả năng sắp xếp đa dạng đường định tuyến trong các giải thuật định tuyến đa dạng như vừa trình bày ở trên, hoạt động cứu định tuyến của mạng trên chip cần được xem xét một cách toàn diện hơn và chi tiết hơn. Đồng thời các nghiên cứu về định tuyến cho mạng trên chip luôn cần phải hướng đến khả năng cân bằng và tối ưu giữa hiệu quả định tuyến và các chi phí thiết kế, thực thi trên vi mạch.

1.4.2 Giải thuật định tuyến tĩnh

Giải thuật định tuyến xác định hay còn gọi là giải thuật định tuyến tĩnh (*static routing*), là các giải thuật xác định cụ thể đường định tuyến của mỗi cặp

nguồn-đích trong cấu trúc liên kết mạng cho trước. Vậy nên, các mạng sử dụng định tuyến tĩnh sẽ luôn lan truyền thông tin theo một đường đi cố định.

Như đã đề cập trong Mục 1.2, định hướng nghiên cứu cấu trúc liên kết dạng lưới hai chiều cho phép thực thi mạng trên chip trong các hệ thống trên chip; khi sắp xếp các thành phần hệ thống trên chip phân bố đều trên phiến silicon. Từ đó, các giải thuật định tuyến cần được phát triển phù hợp theo cả hai tiêu chí là bảo đảm chất lượng dịch vụ - QoS và hiệu quả cân bằng chi phí thiết kế và thực thi trong các ràng buộc của môi trường trên chip. Giải thuật định tuyến ưu tiên theo hướng (DOR: *Dimenssion Order Routing*) là phù hợp với cấu trúc liên kết 2D mesh và các hạn chế của môi trường thực thi trên chip. Các giải thuật DOR là tập hợp các giải thuật có nguyên lý chọn đường định tuyến theo các hướng ưu tiên và chuyển hướng ở góc định tuyến phù hợp. Các giải thuật DOR thực hiện chuyển tin theo hướng ưu tiên được xác lập từ trước, và sẽ cấm một số góc chuyển nhất định nhằm tránh hiện tượng khóa vòng chết.

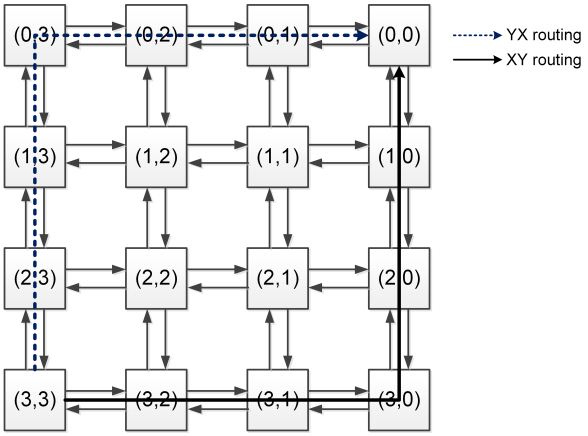


Hình 1.11: Lưu đồ thuật toán định tuyến XY.

Giải pháp định tuyến tĩnh XY [23] là một đại diện điển hình của các giải thuật DOR, rất phù hợp với mạng trên chip do nguyên lý chọn đường định tuyến đơn giản và áp dụng hiệu quả trong các cấu trúc hai chiều (2D mesh hoặc 2D torus). Định tuyến tĩnh XY ưu tiên chuyển thông tin từ nguồn theo phương ngang (X) sau đó chuyển theo phương dọc (Y). Giải thuật chọn đường của định

tuyến tĩnh XY được mô tả như lưu đồ ở Hình 1.11. Trong giải thuật định tuyến XY, khi cặp nguồn-đích có độ lệch phương ngang bằng 0 thì đường định tuyến chỉ theo phương dọc (Y) để hướng đến đích. Khi thay đổi thứ tự ưu tiên của phương truyền tin, thông tin truyền theo phương dọc trước sau đó mới chuyển theo phương ngang ta sẽ có giải thuật định tuyến YX.

Hoạt động chuyển tin trong mạng theo một đường định tuyến nhất định cho phép hệ thống cân bằng tải truyền thông trên các liên kết mạng, đồng thời hoạt động quyết định đường đi và phân xử các yêu cầu trao đổi thông tin đơn giản và hiệu quả. Hạn chế chính của định tuyến tĩnh là khi trạng thái mạng thay đổi sẽ ảnh hưởng lớn đến hoạt động truyền thông do đặc tính không thể thay đổi đường định tuyến phù hợp với trạng thái mạng.



Hình 1.12: Định tuyến tĩnh XY và YX trong mạng 2D-mesh.

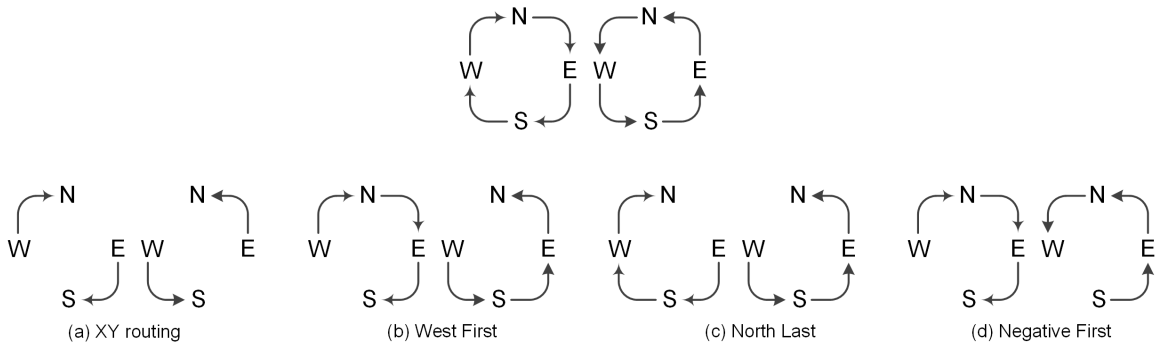
1.4.3 Giải thuật định tuyến thích nghi

Giải thuật định tuyến thích nghi hay còn được gọi là giải thuật định tuyến động (*dynamic routing*), là dạng giải thuật định tuyến linh hoạt trong hoạt động chọn đường định tuyến thông qua các nút mạng của giao tác truyền thông. Định tuyến thích nghi cho phép thay đổi đường định tuyến một cách linh hoạt, thích ứng với sự thay đổi trạng thái mạng trong khi vẫn bảo đảm thông tin đến đích tin cậy. So với các giải pháp định tuyến tĩnh, các giải thuật định tuyến thích nghi có cơ chế phức tạp hơn, việc thực thi bộ định tuyến cũng tốn kém hơn về mặt tài nguyên. Do đó, các giải pháp định tuyến thích nghi cần được nghiên cứu để cân bằng giữa hiệu năng, sự linh hoạt và chi phí thực thi. Sau đây là một số giải thuật định tuyến thích nghi điển hình:

Giải thuật định tuyến góc mở (*Turn model*)

Là một giải thuật định tuyến dạng ưu tiên theo hướng, giải thuật định

tuyến góc mở được Glass và Ni đề xuất năm 1992 cho các mạng máy tính cấu trúc liên kết dạng lưới hai chiều [23]. Dựa trên cấu trúc mạng dạng lưới n chiều và hoạt động chuyển hướng của đường định tuyến nhằm tránh hiện tượng khóa chết (Hình 1.13). Glass và Ni đã chứng minh được rằng cần cấm tối thiểu $n \times (n-1)$ góc rẽ thì giải thuật định tuyến đó sẽ bảo đảm chống được hiện tượng khóa chết trong mạng. Trong cấu trúc mạng dạng lưới hai chiều 2D-mesh, giá trị $n = 2$ và sẽ có 8 góc rẽ khả dĩ, thì một giải thuật định tuyến cần cấm tối thiểu 2 góc rẽ.



Hình 1.13: Góc mở trong giải thuật định tuyến góc mở.

Như trong Hình 1.13, từ 8 góc rẽ và ta cần cấm ít nhất hai góc rẽ sẽ có khá nhiều tổ hợp các góc rẽ được lựa chọn để có các giải thuật định tuyến khác nhau. Dựa vào đặc điểm đối xứng của mạng và thống kế, Glass và Ni đề xuất ba giải thuật định tuyến thích nghi đại diện cho mô hình định tuyến góc rẽ là: định tuyến Phía Tây trước (WF: *West First*), định tuyến Phía Âm trước (NF: *Negative First*) và định tuyến Phía Bắc cuối (NL: *North Last*).

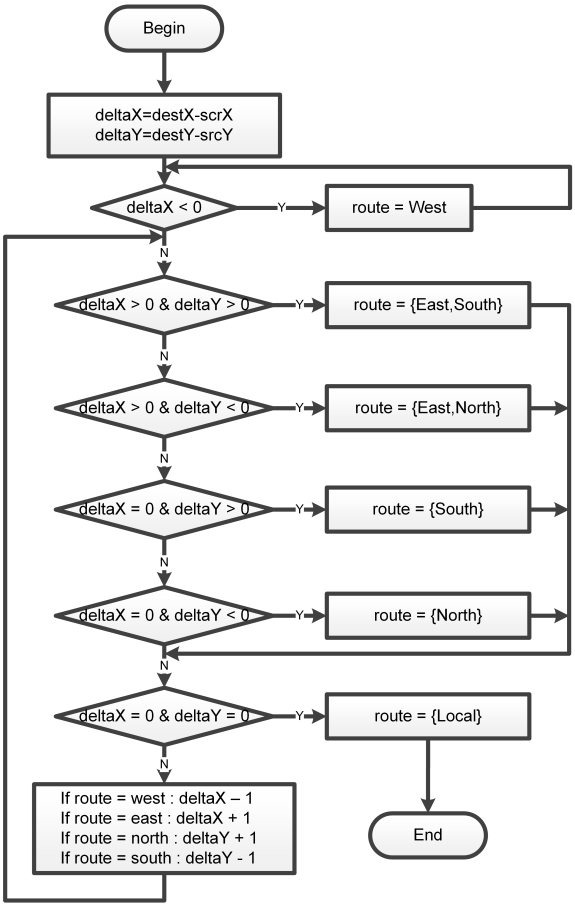
Định tuyến Phía Tây trước

Là giải thuật định tuyến ưu tiên theo hướng (DOR) với nguyên lý cấm hai góc mở về hướng Tây ở Hình 1.13b. Nguyên lý định tuyến được tóm lược như sau:

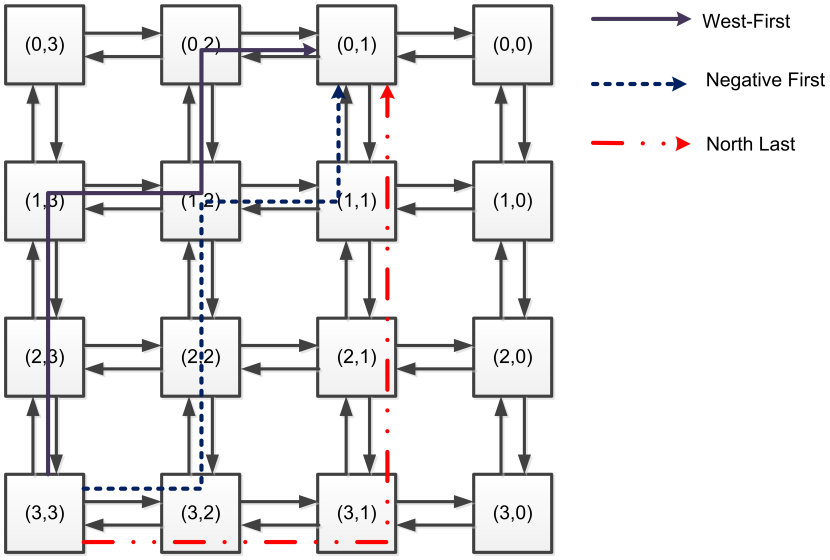
- Thông tin ưu tiên chuyển hướng Tây trước khi chuyển theo các hướng khác.
- Trên đường định tuyến sau khi di chuyển qua hướng Tây, thông tin sẽ được chuyển qua các hướng bất kỳ còn lại.
- Thông tin đang chuyển qua các hướng khác hướng Tây (Bắc, Nam và Đông) sẽ không được chuyển qua hướng Tây. Đây là góc rẽ bị cấm.

Lưu đồ giải thuật định tuyến phía Tây trước được trình bày trong Hình 1.14.

Định tuyến Phía Âm trước



Hình 1.14: Lưu đồ thuật toán định tuyến Phía Tây trước (WF).



Hình 1.15: Các đường định tuyến của nhóm giải thuật Turn model.

Giải thuật định tuyến Phía Âm trước, quy định từ cấu trúc mạng có sự đánh số tăng theo hướng. Hình 1.11 cho thấy theo chiều ngang (X) hướng tăng (*dương*) là từ phải sang trái và theo chiều dọc (Y) hướng tăng là từ trên xuống dưới. Như vậy thì hướng âm sẽ ngược lại với hướng tăng, theo chiều ngang sẽ hướng từ trái sang phải và theo chiều dọc (Y) có hướng từ dưới lên trên. Từ đặc điểm hướng âm trước, cũng có nghĩa là nếu thông tin đang chuyển ở hướng dương sẽ không được rẽ góc về hướng âm. Do đó, trong giải thuật định tuyến này có hai góc rẽ bị cấm là góc rẽ từ Đông sang Bắc và góc rẽ từ Bắc sang Đông (Hình 1.13d).

Định tuyến Phía Bắc cuối

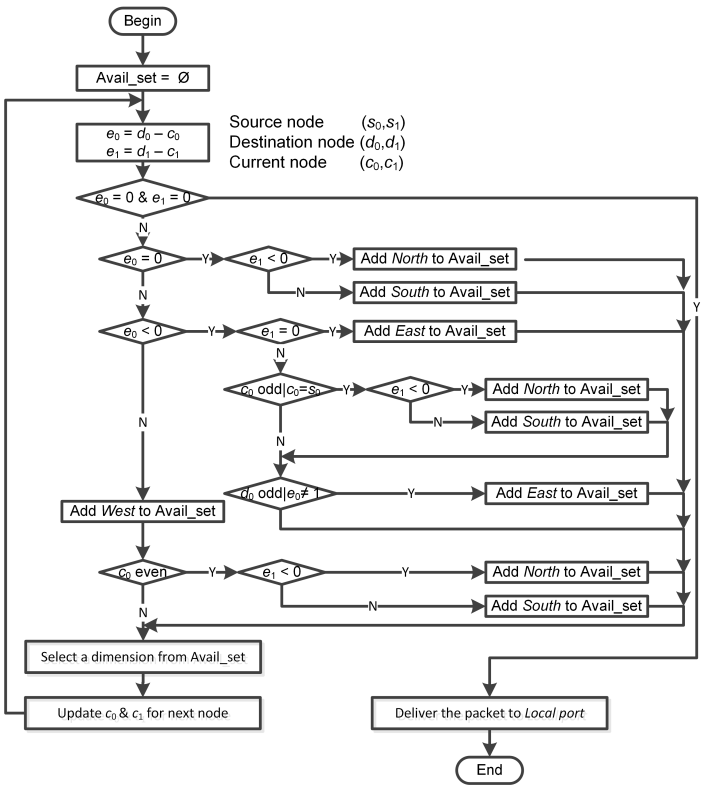
Giải thuật này có cách lựa chọn ưu tiên chuyển hướng khác với hai giải thuật định tuyến trước (WF và NF). Quá trình định tuyến sẽ linh hoạt lựa chọn các hướng chuyển bất kỳ (Đông, Tây và Nam); nhưng khi đã chuyển đến hướng Bắc thì sẽ không chuyển đến các hướng khác mà chỉ hướng đến đích đến của giao tác. Luật cấm góc rẽ của các giải thuật định tuyến WF, NF và NL áp dụng cho cấu trúc hai chiều dạng lưới 2D-mesh được minh họa như trong Hình 1.13.

Hình 1.15 mô tả đường định tuyến được lựa chọn từ ba giải thuật định tuyến ở trên trong cấu trúc mạng lưới 2D-mesh. Do độ lệch giữa nguồn và đích không cho phép hướng thông tin về phía Tây, giải thuật định tuyến WF linh hoạt chọn hai hướng Bắc và Đông để hướng đến đích. Tương tự thì giải thuật định tuyến NF linh hoạt chọn các hướng từ nút nguồn (3,3) đến nút đích (0,1) do cả hai phương đều hướng Âm nên có thể lựa chọn cả hai. Trường hợp của giải thuật định tuyến NL, do có dịch chuyển về Phía Bắc nên thông tin phải được chuyển về hướng Đông hoàn toàn rồi mới chuyển hướng Bắc, vì sau đó chỉ hướng đến đích.

Hoạt động định tuyến của thông tin trong mạng là sự kết hợp của giải thuật định tuyến và các kỹ thuật quyết định định tuyến. Do đó trong một số trường hợp sử dụng giải thuật định tuyến thích nghi nhưng thông tin chỉ hướng theo một tuyến định sẵn từ nguồn tin hoặc theo cấu hình từ trước của mạng.

Định tuyến chẵn-lẻ (OE: *Odd-Even routing*)

Chiu và cộng sự trình bày giải thuật định tuyến thích nghi gọi là giải thuật định tuyến chẵn-lẻ [10]. Giải thuật định tuyến chẵn-lẻ không sử dụng kênh ảo và không cấm chuyển các góc rẽ cố định; thay vào đó được phân chia theo không gian mạng. Mạng được phân chia và các bộ định tuyến được đánh số theo cấu trúc 2D mesh để có các giá trị địa chỉ chẵn/lẻ theo phương ngang (X). Trong quá trình định tuyến, việc chuyển hướng từ X sang Y được thực hiện ở các nút



Hình 1.16: Lưu đồ thực hiện giải thuật định tuyến chẵn-lẻ [10].

mạng chẵn/lẻ phụ thuộc vào hướng đang chuyển là hướng Đông/Tây. Lưu đồ nguyên lý của giải thuật định tuyến chẵn-lẻ được mô tả trong Hình 1.16. Từ lưu đồ cho thấy, giải thuật định tuyến chẵn-lẻ cung cấp tập các hướng truyền khả dĩ mà thông tin dựa vào đó sẽ hướng đến đích nhằm tăng tính linh hoạt cho giải thuật định tuyến. Do vậy, bộ định tuyến cần cơ chế lựa chọn hướng ra tại mỗi điểm trung gian trong tập các hướng khả dĩ phù hợp với trạng thái mạng và đáp ứng tối ưu cho ứng dụng của hệ thống.

1.4.4 Thực hiện định tuyến

Giải thuật định tuyến đề ra nguyên tắc lựa chọn đường đi của thông tin từ nguồn đến đích, do đó cần kết hợp với nguyên tắc tạo nên thông tin định tuyến phù hợp với các điều kiện ràng buộc của hệ thống. Trong mạng máy tính, các bộ định tuyến với tài nguyên đa dạng và khả năng xử lý phong phú thì thực hiện định tuyến rất linh hoạt; hơn nữa tại các bộ định tuyến có thể thực hiện định tuyến với nhiều giải thuật định tuyến phức tạp. Tuy nhiên ở mạng trên chip, hoạt động thực hiện định tuyến cần đơn giản và tối ưu phù hợp với từng hệ thống và cân bằng với chi phí thiết kế và thực thi do phải tuân thủ các ràng buộc trong thiết kế và sản xuất.

Định tuyến phân tán là giải pháp thực hiện định tuyến áp dụng cho

mạng phân tán, khi tất cả các bộ định tuyến đều có chức năng xử lý định tuyến. Định tuyến phân tán cho phép xử lý thông tin định tuyến tại bất kỳ nút mạng nào phù hợp cho các giải thuật định tuyến thích nghi. Đồng thời hoạt động xử lý được trải đều trên toàn mạng sẽ khắc phục được hiện tượng tắc nghẽn trung tâm. Hạn chế của giải pháp này là chi phí thiết kế và thực thi bộ định tuyến lớn, làm gia tăng chi phí thực hiện và chi phí hoạt động của toàn hệ thống lên rất nhiều lần. Đại diện tiêu biểu cho nhóm giải thuật định tuyến này là giải pháp định tuyến phân tán trên cơ sở logic của Flich và nhóm cộng sự [18].

Định tuyến tại nguồn là giải pháp thực hiện định tuyến tại nguồn phát tin, dựa vào giải thuật định tuyến nguồn phát tin quyết định đường định tuyến bằng thông tin định tuyến phù hợp. Thông tin định tuyến sẽ được phát cùng gói tin ở flit tiêu đề. Tại mỗi nút mạng trên đường định tuyến, bộ định tuyến sẽ sử dụng thông tin định tuyến để chỉ định hướng chuyển tin đến bộ định tuyến lân cận [29]. Thực hiện định tuyến tại nguồn sẽ không yêu cầu nhiều xử lý định tuyến tại các bộ định tuyến, và do đó giảm chi phí thực thi bộ định tuyến và giảm trễ truyền thông (các bộ định tuyến chỉ chuyển tiếp thông tin dựa theo thông tin định tuyến đã xác lập tại nguồn phát tin). Nhược điểm của giải pháp định tuyến tại nguồn là thông tin đường định tuyến được xác định tại nguồn tin, đường định tuyến là xác lập cố định và không thể thay đổi ở các nút mạng trung gian. Việc bổ sung trường định tuyến sẽ làm tăng kích thước gói tin đồng thời cũng hạn chế số lượng nút mạng trung gian trên đường định tuyến.

Định tuyến trung tâm sử dụng một khối chức năng định tuyến tập trung cho toàn mạng. Khối chức năng định tuyến này sẽ thực hiện phân xử yêu cầu truyền thông từ ứng dụng và quyết định đường định tuyến cho các giao tác truyền thông trong hệ thống. Giải pháp định tuyến trung tâm có ưu điểm dễ dàng cân bằng tải truyền thông trên toàn mạng; đồng thời cho phép thích ứng với các thay đổi trạng thái của mạng thông qua việc tập hợp thông tin mạng. Vấn đề hạn chế của định tuyến trung tâm là khả năng đáp ứng của khối xử lý trung tâm, dễ xảy ra quá tải khi có quá nhiều yêu cầu truyền thông từ ứng dụng. Manevich và cộng sự đề xuất giải thuật định tuyến thích nghi ATDOR (*Adaptive Toggle Dimension Order Rouing*) là sự kết hợp linh hoạt giữa hai giải thuật định tuyến tĩnh XY và YX thông qua khối điều khiển trung tâm cho mục tiêu cân bằng tải mạng và linh hoạt trong quyết định đường định tuyến dựa trên trạng thái mạng [42].

1.5 Kết luận chương

Mạng trên chip được đề xuất và phát triển đã giải quyết bài toán truyền thông cho các hệ thống trên chip phức tạp một cách hiệu quả, vừa nâng cao được khả năng đáp ứng trao đổi thông tin vừa nâng cao hiệu quả thiết kế. Điều này góp phần dịch chuyển xu thế thiết kế hệ thống sang thiết kế hạ tầng truyền thông, xu thế thiết kế lấy truyền thông làm trung tâm. Việc đưa mô hình mạng cho vào hoạt động truyền thông trên chip vừa thuận lợi khi có rất nhiều kỹ thuật truyền thông đã được biết đến vừa khó khăn khi cần đáp ứng các điều kiện ràng buộc của việc thực thi trên vi mạch.

Chương này đã trình bày các vấn đề cốt lõi của mạng trên chip, từ tổng quan vấn đề truyền thông, mô tả cấu trúc liên kết đến chi tiết các kỹ thuật truyền thông và giải thuật định tuyến được sử dụng để phù hợp với đặc điểm thực thi của mạng trên chip. Trên cơ sở nghiên cứu hoạt động truyền thông mạng trên chip cho thấy cấu trúc liên kết dạng lưới hai chiều 2D-mesh hoặc 2D-Torus là phù hợp với công nghệ chế tạo vi mạch hiện tại. Do vậy, các công trình nghiên cứu mạng trên chip đã và đang được tập trung nghiên cứu sử dụng cấu trúc liên kết này. Công trình [C1] của luận án này cũng được thực hiện nhằm mục đích thực hiện khảo sát chức năng truyền thông của mạng trên chip sử dụng cấu trúc liên kết dạng lưới hai chiều 2D-mesh với các kỹ thuật truyền thông đã trình bày.

Một trong những ưu điểm nổi bật của mạng trên chip là tính mềm dẻo trong truyền thông được đề xuất với nhiều ưu điểm khác nhau và phụ thuộc vào ứng dụng. Tuy vậy, ứng dụng rất đa dạng và luôn thay đổi theo thời gian. Một giải thuật định tuyến hay một cấu trúc liên kết không phải khi nào cũng phát huy lợi thế, đáp ứng được yêu cầu này của ứng dụng. Chưa kể khi hệ thống có độ tích hợp càng cao thì khả năng xảy ra hỏng hóc, lỗi càng lớn. Trước thực tế đó, mạng trên chip cần được nghiên cứu phát triển để có được những khả năng tự thay đổi (hay nói cách khác là tự cấu hình) để nâng cao hiệu quả thiết kế và sử dụng. Xu hướng tái cấu hình dựa trên sự kết hợp nhiều giải thuật định tuyến và cấu trúc liên kết nhằm đáp ứng các thay đổi của thuật tế của ứng dụng là xu thế rất đáng được quan tâm phát triển, và nội dung này sẽ được tập trung đề cập ở chương tiếp theo.

Trên cơ sở nghiên cứu hoạt động truyền thông trên chip theo định hướng mạng trên chip, luận án này đã xây dựng mạng trên chip sử dụng cho cấu trúc 2D-mesh nhưng cũng đồng thời có thể hỗ trợ các cấu trúc dạng lưới khác như

2D-torus hay folded torus [C1]. Trong công trình này đề xuất kiến trúc bộ định tuyến thực hiện chuyển mạch gói kết hợp với kỹ thuật điều chuyển dữ liệu WH và điều khiển luồng credit-based. Hoạt động truyền thông luân chuyển các gói tin được chia nhỏ thành các flit 34 *bit* với 2 *bit* trọng số cao dùng để đánh dấu flit tiêu đề, các flit thân và flit cuối. Một gói tin của giao tác sẽ được đóng gói tại đầu phát với các thông tin điều khiển định tuyến được tính toán và sắp xếp vào trường định tuyến (PTT:*Path-To-Target*) gồm 18 *bit* trọng số thấp của flit tiêu đề. Bộ định tuyến xử lý thông tin định tuyến dựa vào 2 *bit* thấp của trường định tuyến để chuyển hướng đến lối ra được lựa chọn từ trước, và sau đó sẽ xóa 2 *bit* này để chuyển tiếp đến vị trí tiếp sau.

Chương 2

Vấn đề tái cấu hình và truyền thông tái cấu hình

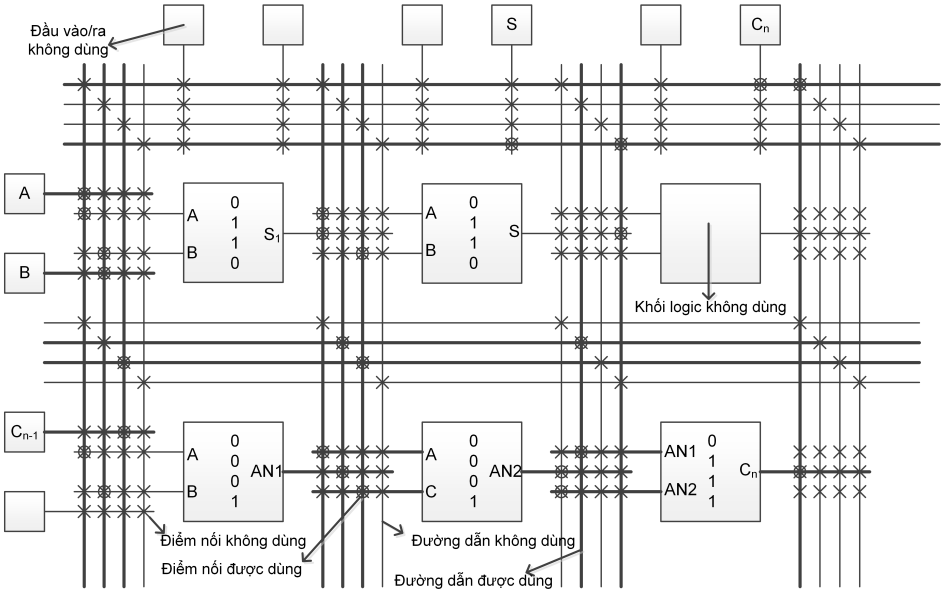
Trong xu thế phát triển nhanh và mạnh của thiết kế hệ thống, thị trường mua bán lõi IP phát triển ngày một đa dạng cho phép người thiết kế giảm thiểu thời gian và chi phí phát triển từng chức năng của ứng dụng. Lúc này thiết kế hệ thống chỉ cần tập trung vào quá trình tổ chức hoạt động của hệ thống và liên kết các lõi chức năng một cách tối ưu phù hợp với từng ứng dụng, và cũng từ đó hình thành xu thế thiết kế lấy truyền thông làm trung tâm. Vì dựa trên khả năng tái sử dụng các lõi chức năng đã có, hệ thống hoạt động cần được phát triển các khả năng sắp xếp linh hoạt các thành phần chức năng một cách hợp lý. Trong hoạt động sắp xếp các thành phần cấu thành, khả năng tự thích ứng các thay đổi cấu hình bổ sung một tính năng rất quan trọng chính là khái niệm tái cấu hình.

Đáp ứng nhu cầu phát triển nhanh và rộng của các hệ thống nhúng trong nhiều lĩnh vực, các thiết kế hệ thống trên chip cần được nghiên cứu để bổ sung nhiều tính năng hơn cho các sản phẩm nhằm nâng cao hiệu quả hoạt động nhưng tiết giảm các nhiều chi phí. Xu thế thiết kế tái cấu hình cung cấp khả năng tự thích ứng cho phép tái sử dụng thiết kế trong nhiều ứng dụng, đồng thời nâng cao hiệu quả hoạt động của hệ thống bằng cách tự thay đổi để thích ứng với nhiều điều kiện hoạt động khác nhau. Hơn nữa, trong sự phát triển của các hệ thống trên chip với công nghệ sản xuất tiên tiến đã cho phép tích hợp rất nhiều thành phần thì khả năng sai hỏng trong sản xuất và hoạt động là rất lớn. Các thiết kế tái cấu hình khi được kết hợp với khả năng phát hiện sai hỏng (BIST: *Built-In Self Test*) cho phép hệ thống khắc phục được một số khiếm khuyết mà vẫn hoạt động ổn định và hiệu quả [24, 32].

2.1 Hệ thống trên chip và định hướng tái cấu hình

Các hệ thống trên chip cần được thiết kế với độ linh hoạt cao nhằm đáp ứng các yêu cầu đa dạng của ứng dụng [21], đồng thời để nâng cao hiệu quả thiết kế, tiết giảm các chi phí nghiên cứu lần đầu (NRE: *Non Recurring Engineering cost*) và thời gian đưa sản phẩm ra thị trường. Hiện tại, thiết kế hệ thống được phát triển theo hai định hướng chính: sử dụng các bộ vi xử lý đa dụng (GPP: *General Purpose Processor*) kết hợp với các thành phần ngoại vi; hoặc sử dụng các mạch tích hợp chuyên dụng (ASIC: *Application Specific Integrated Circuit*).

Hướng đến khả năng thích ứng cho dải rộng ứng dụng, các hệ thống cần sự linh hoạt trong quá trình tính toán và xử lý dựa trên tập các tài nguyên phần cứng và phần mềm. Các hệ thống sử dụng vi xử lý đa dụng có khả năng linh hoạt cho rất nhiều tác vụ, tạo nên thiết bị đa năng như máy tính. Tuy nhiên, các hệ thống này tiêu thụ năng lượng lớn và phụ thuộc nhiều vào khả năng lập trình. Vậy nên, một hệ thống có khả năng linh hoạt dựa trên khả năng tái sắp xếp các tài nguyên phần cứng phù hợp với ứng dụng đa dạng là một xu thế tối ưu cần được quan tâm nghiên cứu. Hệ thống được thiết kế theo xu hướng có khả năng tái tổ chức và sắp xếp các thành phần được gọi chung là hệ thống tái cấu hình (RSoC: *Reconfigurable System-on-Chip*).



Hình 2.1: Mạch cộng tái cấu hình bằng FPGA.

Các thiết kế sử dụng công nghệ FPGA là minh họa rõ nét và phổ biến của khả năng tái cấu hình thông qua việc thiết lập các kết nối và chức năng của các khối logic. Hình 2.1 minh họa cho khả năng tái cấu hình để thực thi chức năng mạch cộng bằng FPGA. Thông qua khả năng lập trình sẽ cho phép thiết lập

các khối logic thành các hàm logic cụ thể để thực hiện các chức năng phù hợp ứng dụng và sử dụng công cụ phù hợp để cấu hình các đường dẫn tạo các liên kết giữa các khối được chọn. Do vậy, các thiết kế tái cấu hình cho phép thay đổi cách tổ chức sắp xếp các tài nguyên của mình để thực thi các chức năng logic của hệ thống. Hình 2.1 sử dụng khả năng tái cấu hình của FPGA để thực hiện hai hàm logic S và C_n.

$$S = A.B.\bar{C}_{n-1} + A.\bar{B}.C_{n-1} + \bar{A}.B.C_{n-1} = A \oplus B \oplus C_{n-1} \tag{2.1}$$

$$C_n = A.B + (A \oplus B)C_{n-1} \tag{2.2}$$

Công nghệ FPGA cho phép thực thi linh hoạt các hệ thống thống tận dụng tối đa khả năng thay đổi cấu hình, để hướng đến các hệ thống tái cấu hình động từng phần (*dynamic partial reconfigurable*) [31, 55]. Hoạt động thay đổi cấu hình dựa trên công nghệ FPGA sẽ không áp dụng được cho các hệ thống thực thi trên ASIC, vì sử dụng giải pháp tái sắp xếp tài nguyên của FPGA dựa trên đặc điểm khả trình của công nghệ FPGA. Đồng thời, xu thế tái cấu hình được triển khai cho các hệ thống được thực hiện đầy đủ từ đầu nhưng vẫn có khả năng tương thích cho nhiều trường hợp thay đổi cấu hình (do sai hỏng trong sản xuất, yêu cầu ứng dụng,...) được thực hiện trên cả công nghệ ASIC và FPGA. Trong luận án này, khái niệm tái cấu hình dùng để chỉ các hướng nghiên cứu về các hệ thống có khả năng tự thích ứng với các thay đổi cấu hình từ nhu cầu khách quan để nâng cao tính hiệu quả của thiết kế.

Hệ thống trên chip (SoC) được sử dụng trong các hệ nhúng cho các thiết bị thông minh cần khả năng đáp ứng yêu cầu tương tác người dùng và thích nghi với điều kiện hoạt động. Điều này cần các thiết kế có khả năng thay đổi hoạt động một cách thông qua các chương trình phần mềm chạy trên các vi xử lý đa dụng hoặc vi xử lý có tập lệnh chuyên dụng (ASIP: *Application Specific Instruction set processor*). Do đó, hiện tồn tại hai hướng thiết kế hệ thống chính cho các hệ thống nhúng là:

- Hệ thống hoạt động theo nguyên lý luồng dữ liệu (*data flow-driven*) với đặc điểm sử dụng các khối logic trực tiếp xử lý các luồng dữ liệu để thực hiện các chức năng của ứng dụng cho phép hệ thống hoạt động tối ưu và ổn định. Các hệ thống được thiết kế cho các ứng dụng cụ thể luôn được tối ưu cho các hoạt động đã biết nên có hiệu quả cao, tiêu hao năng lượng thấp và có giá thành rẻ khi sản xuất với số lượng lớn. Hạn chế lớn nhất của hướng thiết kế ASIC là giới hạn dải tùy biến ứng dụng và thời gian

đưa sản phẩm ra thị trường kéo dài do cần nhiều thời gian cho nghiên cứu và đánh giá toàn diện [21].

- Hệ thống hoạt động theo tập lệnh (*instruction-driven*) cho phép hoạt động tùy biến linh hoạt dựa trên khả năng lập trình. Hoạt động của hệ thống có khả năng thích ứng trong dải rộng các ứng dụng bằng cách thay đổi chuỗi lệnh thực thi trên các vi xử lý đa dụng. Điểm hạn chế của hệ thống dựa trên vi xử lý là mức độ tiêu thụ năng lượng cao, không phù hợp với các thiết bị sử dụng nguồn năng lượng giới hạn.

Tập trung cho định hướng tái sắp xếp các thành phần bên trong hệ thống, hệ thống tái cấu hình góp phần tối ưu cho dải rộng các ứng dụng, nâng cao hiệu quả thiết kế hệ thống và giảm thiểu các chi phí (NRE, thời gian đưa ra thị trường,...). Dựa trên đặc điểm của các phần tử tái cấu trong các tác vụ tính toán và xử lý, ta có thể phân các hệ thống tái cấu hình thành ba loại sau [21, 69]:

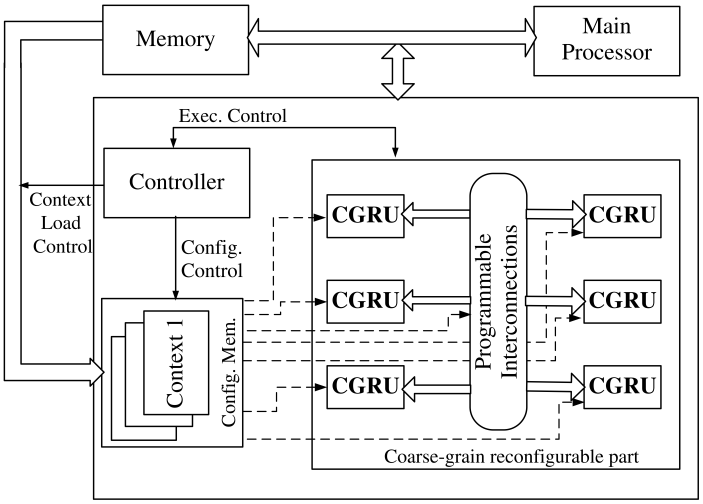
- Kiến trúc tái cấu hình hạt nhân tinh (FGRA: *Fine-Grained Reconfigurable Architecture*) là những hệ thống có khả năng tái cấu hình linh hoạt dựa trên các thành phần và kết nối có thể cấu hình chi tiết mức bit. Tuy nhiên, do phải cấu hình mức ở chi tiết, các hệ thống này có hạn chế là mức tiêu thụ năng lượng cao và hoạt động cấu hình khá phức tạp (cần nhiều thông tin cấu hình để thiết lập hệ thống)[22].

- Kiến trúc tái cấu hình hạt nhân thô (CGRA: *Coarse-Grained Reconfigurable Architecture*) là kiến trúc cho phép tái cấu hình dựa trên các thành phần tái cấu hình là các khối logic và kết nối hoạt động ở mức từ (*word-level*). Do sử dụng cơ chế hoạt động với các hạt nhân thô, hệ thống hoạt động xử lý và luồng dữ liệu theo các đơn vị tin (từ) cho phép nâng cao hiệu quả hoạt động và giảm thiểu tiêu đề tái cấu hình để giảm công suất tiêu thụ và không gian thực thi [65].

- Kiến trúc tái cấu hình hạt nhân trung (MGRA: *Middle-Grained Reconfigurable Architecture*) là định dạng kiến trúc tái cấu hình hoạt động với các đơn vị tin tương đối nhỏ (dưới $8bit$). Giải pháp kiến trúc MGRA được đề xuất nhằm tận dụng ưu điểm của hai loại kiến trúc trên, cân bằng khả năng sắp xếp linh hoạt các thành phần và chi phí hoạt động (công suất, tiêu đề tái cấu hình, ...) của hệ thống.

Về cơ bản một hệ thống tái cấu hình có hạt nhân thô (CGRS) bao gồm các thành phần cơ bản như trong Hình 2.2 gồm khối điều khiển (***Controller***),

bộ nhớ ngữ cảnh (***Context***), và các đơn vị tái cấu hình hạt nhân thô (**CRGU: *Coarse-grain Reconfigurable Unit***) được kết nối với nhau một cách linh hoạt bằng các liên kết khả trình [66]. Trong đó khối ***Controller*** có chức năng điều khiển mọi hoạt động cấu hình (*thu nhận trạng thái cấu hình, thiết lập cấu hình, ...*); khối ***Context*** là bộ nhớ để lưu giữ thông tin cấu hình của tập hợp các đơn vị cấu hình; và các đơn vị tái cấu hình là các khối chức năng logic đã hoàn thiện cho từng hoạt động logic cụ thể của hệ thống. Kiến trúc sử dụng khối điều khiển và bộ nhớ ngữ cảnh cung cấp khả năng thay đổi nội dung thông tin cấu hình của hệ thống một cách linh hoạt động ngay cả khi hệ thống đang hoạt động là cơ sở cho phép hệ thống có thể tái cấu hình ngay khi đang chạy (*run-time*).



Hình 2.2: Kiến trúc cơ bản hệ thống tái cấu hình hạt nhân thô.

Với đặc điểm linh hoạt khả năng tính toán và xử lý trong các hệ thống trên chip, hệ thống tái cấu hình trên cơ sở cấu trúc lõi thô hiện đã có nhiều nghiên cứu và đề xuất các giải pháp rất đa dạng đến từ nhiều trung tâm nghiên cứu ở nhiều nước khác nhau. Các kiến trúc tái hệ thống tái cấu hình được phát triển khá đa dạng đáp ứng cho nhiều định hướng ứng dụng và mục tiêu nghiên cứu. Nhưng trên cơ sở các hạt nhân thô thường là các lõi tính toán và xử lý logic, các kiến trúc hệ thống tái cấu hình thường hướng đến khả năng xử lý tín hiệu hoặc tính toán phức hợp như: RaPiD [16], MorphoSys [60]. Cùng trong xu thế phát triển của các thiết bị thông minh và đa phương tiện các hệ thống tái cấu hình cho phép tối ưu chi phí và nâng cao hiệu quả hoạt động của hệ thống như trong ADRES [44] từ trung tâm IMEC của Vương quốc Bỉ, và REMUS-II [28] từ Đại học Đông Nam, Trung Quốc. Một số kiến trúc phức tạp hơn để hướng đến nhiều ứng dụng hơn được phát triển bằng sự kết hợp nhiều kiến trúc và công cụ phát triển như: XPP-III [64] từ PACT/XPP Technologies của Đức, MORPHEUS [69]

của Châu Âu. Mỗi kiến trúc được đề xuất đều có những đặc điệm cụ thể riêng và hướng đến một nhóm các ứng dụng cụ thể (*xử lý tín hiệu, tính toán phức hợp, đa phương tiện*, ...) để tối ưu kiến trúc phù hợp.

Giải pháp tái cấu hình cho hệ thống dựa trên các đơn vị chức năng linh hoạt đáp ứng tối ưu cho dải rộng các ứng dụng đã khẳng định hướng phát triển của lĩnh vực thiết kế tích hợp. Tuy nhiên, các hệ thống tái cấu hình chủ yếu sử dụng khối tái cấu hình có sẵn dạng lưới hai chiều kết hợp với thông tin cấu hình tập trung để sắp xếp các thành phần để thích ứng với hoạt động ứng dụng. Do vậy, với các hệ thống hoạt động phức hợp có nhiều thành phần và có nhiều yêu cầu thay đổi cấu hình sẽ ảnh hưởng lớn đến hiệu năng của khối điều khiển trung tâm. Như trong giải pháp kiến trúc MORPHEUS sử dụng STNoC để kết nối các thành phần tái cấu hình thì trong kiến trúc XPP-III [64] sử dụng mạng trên chip cho phép nâng cao hiệu quả truyền thông ngay trong thành phần tái cấu hình cũng là một giải pháp góp phần tối ưu toàn diện hơn cho các hệ thống tái cấu hình. Nhưng để thích ứng với sự linh hoạt bố trí sắp xếp các thành phần của hệ thống tái cấu hình, hoạt động truyền thông của hệ thống cần được giải quyết một cách toàn diện hơn và tối ưu hơn trong điều kiện tái cấu hình. Một giải pháp truyền thông kết hợp được khả năng truyền thông băng thông cao và tự thích nghi với các thay đổi cấu hình của hệ thống. Điều này sẽ góp phần quan trọng trong các thiết kế tái cấu hình, nhằm đem lại những sản phẩm đáp ứng tối ưu cho nhiều ứng dụng cũng như nâng cao hiệu quả thiết kế cho phép tái thích ứng lỗi sai.

2.2 Vấn đề tái cấu hình đối với mạng trên chip

Nhằm đáp ứng các nhu cầu phát triển ứng dụng, thiết kế hệ thống trên chip luôn cần được nghiên cứu và phát triển để tạo ra nhiều sản phẩm tốt hơn, nhiều thiết kế hiệu quả hơn. Từ nhu cầu phát triển của các ứng dụng đa phương tiện và thiết bị thông minh thì hệ thống trên chip được phát triển với nhiều chức năng mới phức hợp trên một chip đơn. Do đó, giải pháp mạng trên chip ra đời và phát triển đã và đang góp phần giải quyết một cách toàn diện và hiệu quả các vấn đề truyền thông trên chip cũng như khả năng nâng cao hiệu quả hoạt động của toàn hệ thống. Ngoài ra, như trong Mục 2.1 vừa đề cập thì xu thế phát triển của các thiết kế tái cấu hình góp phần quan trọng để nâng cao hiệu quả thiết kế và thực thi của các hệ thống trên chip trong những năm gần đây. Do vậy, định hướng kết hợp cả hai thách thức trên trong các thiết kế hệ thống

trên chip cần được nghiên cứu để giải quyết một cách toàn diện những vấn đề đặt ra trong sự phát triển của lĩnh vực thiết kế. Mạng trên chip đã và đang được phát triển với rất nhiều kết quả khả quan được dùng cho nhiều hệ thống đa chức năng với các hoạt động phức hợp, nhưng vẫn cần được quan tâm nghiên cứu nhiều hơn nữa để đáp ứng tốt hơn những ứng dụng phức hợp đang cần được xây dựng [43]. Mạng trên chip tái cấu hình (RNoC: *Reconfigurable Network-on-Chip*) được hình thành và phát triển trên cơ sở kết hợp các thế mạnh của thiết kế tái cấu hình và giải pháp truyền thông sử dụng mô hình mạng linh hoạt. Giải pháp mạng trên chip tái cấu hình cho phép kết hợp các lõi IP đa dạng trên nền một hạ tầng truyền thông linh hoạt, phù hợp với xu thế thiết kế lấy truyền thông làm trung tâm đang rất được quan tâm hiện nay. Thông qua những vấn đề nổi bật của hệ thống tái cấu hình đã được xem xét ở Mục 2.1, trong phần này sẽ tập trung cho vấn đề truyền thông linh hoạt trên cơ sở sử dụng mô hình mạng để làm rõ hơn các vấn đề cần quan tâm của những giải pháp thiết kế mạng trên chip tái cấu hình.

2.2.1 Động lực và thách thức của mạng trên chip tái cấu hình

Hiện tại, các hệ thống trên chip ngày càng được phát triển đa dạng để đáp ứng nhiều yêu cầu ứng dụng như các hệ thống tái cấu hình được đề cập trong Mục 2.1. Cùng với đó là sự phát triển của xu thế ứng dụng các thiết kế lõi chức năng có sẵn từ thị trường cấp phép lõi IP đã và đang hướng đến xu thế thiết kế hệ thống lấy truyền thông làm trung tâm. Trong xu thế này, người thiết kế phát triển các hệ thống trên chip với ít chi phí hơn và thời gian đưa sản phẩm ra thị trường nhanh hơn để tạo tính cạnh tranh cao hơn. Nhưng để kết hợp được nhiều lõi chức năng được phát triển độc lập thì giải pháp truyền thông trên chip cần được thiết kế một cách linh hoạt và hiệu quả để đảm bảo hệ thống hoạt động ổn định. Hoạt động truyền thông trên chip sử dụng mô hình mạng trên chip đang được xem làm một giải pháp hiệu quả và toàn diện, đáp ứng cho nhiều hệ thống phức hợp [1]. Để đáp ứng tối ưu hơn trong định hướng thiết kế tái cấu hình, các giải pháp tái cấu hình mạng trên chip đã và đang được quan tâm nghiên cứu nhằm giải quyết một cách tối ưu và hoàn thiện hơn theo các mục tiêu sau:

- **Đáp ứng tối ưu cho nhiều ứng dụng**
  Hoạt động truyền thông của hệ thống linh hoạt để đáp ứng yêu cầu trao đổi dữ liệu giữa các thành phần, từ đó nâng cao hiệu quả hoạt động tổng thể của hệ thống [64, 69]. Khi các thành phần chức năng trong hệ thống thay đổi thì nhu cầu trao đổi dữ liệu giữa chúng cũng thay đổi theo. Khi

đó, kiến trúc truyền thông của hệ thống cần được tái cấu hình để đảm bảo cung cấp khả năng truyền thông tối ưu nhất (hiệu năng, công suất tiêu thụ,...). Ví dụ, một số giải pháp mạng tái cấu hình với khả năng tái sắp xếp các thành phần truyền thông được đề xuất giải pháp hiệu quả [7, 31] nhằm bảo đảm chất lượng đáp ứng dịch vụ - QoS và giảm các chi phí hoạt động.

- **Nâng cao hiệu quả thiết kế**
  Khi mật độ tích hợp của hệ thống trên chip càng lớn thì khả năng xảy ra sai hỏng trong quá trình sản xuất càng lớn. Trong nhiều trường hợp, sai hỏng chỉ tập trung ở một số khối chức năng trong hệ thống và các thành phần phần chức năng khác vẫn có thể hoạt động bình thường. Hệ thống vẫn có thể đáp ứng được phần nào yêu cầu của ứng dụng và vẫn có giá trị sử dụng. Nhiều nghiên cứu với mục tiêu phát hiện sai hỏng cũng như giải pháp thiết kế chống sai hỏng được phát triển và đã thu được rất nhiều kết quả khả quan [24, 32, 73]. Một khi phát hiện sai hỏng, việc đưa ra các kiến trúc có khả năng thích ứng với sai hỏng để bảo đảm hệ thống vẫn hoạt động (dù ở chế độ hạn chế) sẽ góp phần tiết kiệm chi phí và nâng cao hiệu quả sử dụng của sản phẩm.

- **Tối ưu hoạt động truyền thông**
  Trong điều kiện hoạt động đa dạng của các hệ thống phức hợp trong các thiết bị thông minh, các thành phần chức năng của hệ thống linh hoạt thực hiện các hoạt động tính toán và đặt ra các yêu cầu truyền thông ngẫu nhiên. Do đó, hoạt động truyền thông trong toàn hệ thống luôn thay đổi giữa các thành phần lõi chức năng và theo thời gian. Giải pháp tái cấu hình truyền thông cung cấp khả năng thích nghi theo các điều kiện hoạt động của hệ thống để tối ưu đáp ứng cho các giao tác truyền thông trong các hệ thống phức hợp; góp phần nâng cao hiệu quả hoạt động của hệ thống trên chip. Theo định hướng này các giải pháp mạng trên chip tái cấu hình được phát triển theo hướng thích ứng với khả năng tái sắp xếp các thành phần truyền thông tối ưu theo ứng dụng như kiến trúc mạng trên chip tái cấu trúc liên kết mạng được đề xuất trong [61, 62].

Mạng trên chip tái cấu hình - RNoC là thuật ngữ dùng để chỉ các giải pháp thiết kế truyền thông trên chip đáp ứng cho các hệ thống trên chip có khả năng thích ứng với sự thay đổi cấu hình mạng mà vẫn bảo đảm chức năng truyền thông cho các lõi IP cho phép hệ thống hoạt động ổn định.

Nhờ khả năng tự thích ứng với sự thay đổi cấu hình, mạng trên chip tái cấu hình là giải pháp truyền thông đáp ứng linh hoạt cho các hệ thống phức hợp với nhiều ưu điểm nổi trội. Cũng như giải pháp mạng trên chip đang là định hướng thu hút sự quan tâm nghiên cứu, giải pháp tái cấu hình mạng trên chip cần được nghiên cứu đánh giá một cách toàn diện trong điều kiện ràng buộc của việc thực thi trên chip. Hướng đến một giải pháp truyền thông hiệu quả, mạng trên chip cần những giao thức truyền thông đủ mạnh và tài nguyên truyền thông lớn, nhưng điều này lại là hạn chế đối với hệ thống trên chip.

Hoạt động truyền thông của các hệ thống trên chip cần đáp ứng tối ưu cho các yêu cầu trao đổi dữ liệu với khối lượng lớn, tốc độ cao và yêu cầu cao về độ tin cậy. Ưu điểm của mạng trên chip là khả năng cung cấp các kết nối linh hoạt tốc độ cao cho các hoạt động truyền thông *end-to-end*, nhưng vẫn tồn tại nhược điểm về khả năng điều khiển cân bằng hoạt động truyền thông do nguyên lý xử lý phân tán so với khả năng phân xử trung tâm của BUS chia sẻ. Vậy nên các giải pháp tái cấu hình khi cung cấp khả năng tự thích ứng với cấu hình mạng cần phải được nghiên cứu đánh giá khả năng bảo đảm chất lượng dịch vụ truyền thông trong toàn hệ thống trong mọi trường hợp thay đổi cấu hình.

Trên cơ sở là các phần tử và mạch điện tích hợp để thực hiện các giao thức truyền thông, mạng trên chip cần những giao thức linh hoạt nhưng phải phù hợp với khả năng thực thi trên chip. Do đó, các nguyên tắc truyền thông cần được mô tả tường minh bởi các trạng thái mạch logic hoặc máy trạng thái hữu hạn. Và trong điều kiện thực thi trên bán dẫn thì các giao thức càng phức tạp sẽ cần càng nhiều thành phần mạch điện để thực hiện nên sẽ yêu cầu càng nhiều tài nguyên của phiến bán dẫn (chi phí không gian thực thi). Vậy nên, cũng như tất cả các thiết kế trên chip các giải pháp tái cấu hình mạng trên chip cần được phát triển trên cơ sở cân bằng giữa hiệu quả linh hoạt truyền thông và chi phí thực hiện trên vi mạch. Thách thức này đặt ra cho các thiết kế cần tối ưu các kiến trúc để thực hiện giải pháp nhằm đáp ứng tốt mục tiêu đặt ra của bài toán thiết kế.

Song song với chi phí thiết kế và thực thi hệ thống, các chi phí hoạt động của hệ thống ngày càng được quan tâm và nhất là đối với các thiết bị thông tin cầm tay có nguồn năng lượng giới hạn. Các giải pháp mạng trên chip sử dụng kiến trúc truyền thông dạng ống (*pipeline*) cùng nguyên lý xử lý phân tán sẽ có tác động lớn đến sự tiêu hao năng lượng của hệ thống. Một giải pháp tái cấu hình hoạt động với các kiến trúc phức tạp ngoài việc gia tăng chi phí thiết kế thực thi cũng sẽ làm gia tăng tiêu hao năng lượng và làm giảm hiệu quả hoạt

động. Nhưng đồng thời các giải pháp tái cấu hình cũng được hướng đến khả năng tối ưu hoạt động động truyền thông cho mục tiêu tiết giảm chi phí hoạt động truyền thông trong hệ thống.

2.2.2 Phân loại giải pháp tái cấu hình mạng trên chip

Mạng trên chip đáp ứng yêu cầu truyền thông cho các hệ thống phức hợp cần thể hiện đầy đủ các hoạt động đặc trưng của môi trường mạng và những ràng buộc do thực thi trên chip. Giải pháp tái cấu hình mạng trên chip đã và đang thu hút sự quan tâm nghiên cứu của nhiều nhóm tác giả theo nhiều mục tiêu và định hướng đa dạng. Vậy nên việc tổng hợp và sắp xếp phân loại các công trình liên quan để làm cơ sở nghiên cứu chuyên sâu là rất cần thiết, đồng thời hoạt động này cũng sẽ làm rõ định hướng nghiên cứu và trọng tâm của hoạt động tái cấu hình mạng trên chip hiện tại. Do vậy, trong phần này sẽ tập trung vào phân tích các cơ sở phân loại giải pháp tái cấu hình mạng trên chip.

Trên cơ sở mục tiêu của các giải pháp tái cấu hình mạng trên chip, chúng ta có thể phân chia các giải pháp mạng trên chip tái cấu hình thành các loại như sau:

- ***Thích ứng cho nhiều ứng dụng***
  Hướng đến mục tiêu hệ thống tự thay đổi để thích ứng với một dải rộng các ứng dụng, giải pháp tái cấu hình mạng trên chip cung cấp khả năng truyền thông thích nghi với sự thay đổi cấu hình hệ thống. Trên đặc điểm của công nghệ FPGA cho phép thiết lập các cấu hình một cách linh hoạt các thiết kế, các giải pháp mạng trên chip tái cấu hình cũng được phát triển để phát huy tối đa khả năng tự thích nghi của thiết kế đối với các ứng dụng. Do vậy, giải pháp tái cấu hình mạng trên chip dựa trên công nghệ FPGA đã và đang được nghiên cứu hướng đến khả năng thích ứng với nhiều thay đổi cấu trúc liên kết từ dạng lưới hai chiều 2D-mesh có những công trình tiêu biểu như: CuNoC của Jovanovic và các cộng sự [31], CoNoChi [54, 55].

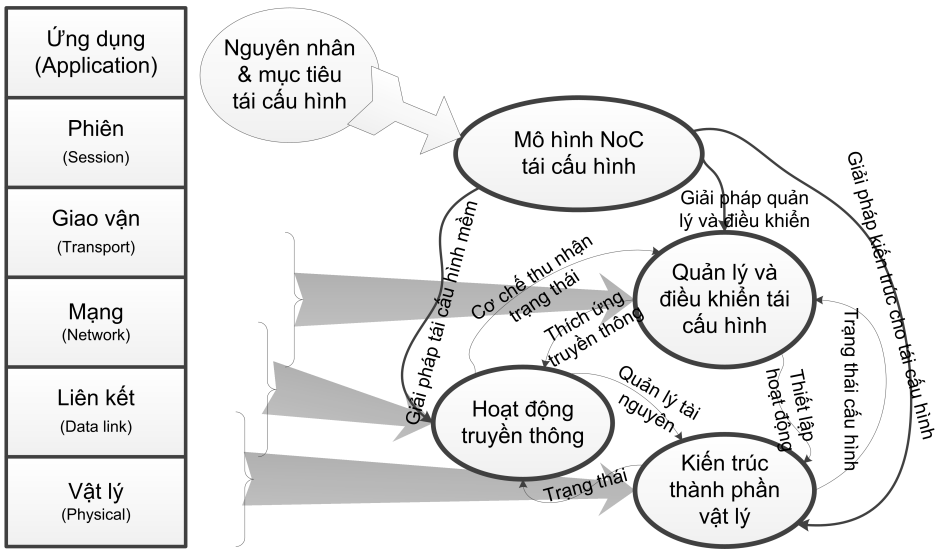
- ***Khả năng chống sai hỏng***
  Trong xu thế phát triển của các thiết kế phức hợp thì nghiên cứu phát hiện sai hỏng và chống sai hỏng đang rất được quan tâm, và mạng trên chip cũng được phát triển nhiều theo định hướng đó. Với đặc điểm khối hóa các thành phần của kiến trúc truyền thông của mạng trên chip hiện đã có nhiều công trình nghiên cứu khả năng kiểm tra phát hiện sai hỏng

(DFT: *Design-for-Test*); điển hình là công trình [68]. Một định hướng cũng đang được quan tâm phát triển của các giải pháp mạng trên chip tái cấu hình là khả năng chống sai hỏng (*Fault-tolerant*) dựa trên khả năng thích nghi với điều kiện cấu hình thay đổi do sai hỏng đã được phát hiện. Các giải pháp tái cấu hình chống sai hỏng được phát triển rất đa dạng; từ khả năng quản lý cấu hình của hệ điều hành để thích ứng với sự thay đổi [71]; hoặc sử dụng các giải pháp định tuyến chống sai hỏng [45, 73]. Tuy nhiên, giải pháp hoàn thiện cho cả hai mục tiêu phát hiện và chống sai hỏng hiện đang còn chưa được quan tâm phát triển để có được những thiết kế mạng trên chip hiệu quả trong thực tiễn hiện nay.

- ***Nâng cao hiệu quả truyền thông và cân bằng chi phí***
  Khả năng tái cấu hình của truyền thông mạng trên chip được phát triển cho phép đáp ứng một cách hiệu quả hơn các yêu cầu trao đổi thông tin của ứng dụng trong các hệ thống trên chip. Bằng khả năng tái sắp xếp các cấu trúc liên kết trong quá trình thiết lập cho phép kiến trúc ReNoC của nhóm tác giả Stensgaard và cộng sự tối ưu hạ tầng truyền thông mạng trên chip với nhiều ứng dụng cụ thể [61, 62]. Cùng trong mục tiêu này, các giải pháp sử dụng khả năng định tuyến tự thích ứng với trạng thái hoạt động của hệ thống sẽ cho phép giải quyết các yêu cầu truyền thông đột biến trong quá trình hệ thống đang hoạt động, như giải thuật định tuyến trung tâm của Manevich và cộng sự [42], hay dựa vào khả năng linh hoạt băng thông giữa các bộ định tuyến của kiến trúc BiNoC [33].

Với đặc thù của thiết kế chip, mỗi sản phẩm là một hệ thống hoàn thiện cho phép hoạt động hiệu quả với những ứng dụng cụ thể. Khả năng tự thích nghi với các thay đổi cấu hình cần được xem xét trong một mô hình tổng thể với những tương quan chức năng của những hoạt động thay đổi cụ thể. Nhằm cung cấp một cái nhìn toàn cảnh các giải pháp tái cấu hình, mạng trên chip tái cấu hình cần được tham chiếu với mô hình phân lớp chức năng của các hoạt động truyền thông [26, 12].

Dựa trên nguyên tắc phân chia chức năng của mô hình phân lớp và nguyên lý hoạt động của tái cấu hình như mô tả trong Hình 2.3, chúng ta có thể phân loại các giải pháp tái cấu hình mạng trên chip dựa theo phương thức hoạt động truyền thông. Dù hoạt động truyền thông được thể hiện qua tất cả các lớp chức năng, các hoạt động truyền thông mạng trên chip chủ yếu tập trung ở các lớp chức năng hướng mạng (ba lớp dưới như đã đề cập ở Chương 1). Tương ứng với hoạt động truyền thông, các giải pháp tái cấu hình mạng trên chip được phân



Hình 2.3: Hoạt động tái cấu hình dành cho mạng trên chip.

chia dựa theo phương thức hoạt động truyền thông theo ba nhóm chức năng như sau:

1. ***Hoạt động quản lý cấu hình và điều khiển truyền thông***
   Xét trong điều kiện cấu hình thay đổi và để đảm bảo hệ thống hoạt động ổn định cần có hoạt động tổng hợp thông tin và quản lý cấu hình ở mức cao. Với đặc điểm hoạt động của mạng trên chip, khả năng quản lý cấu hình sẽ hình thành ở chức năng của lớp mạng bởi chức năng giao tiếp mạng (NI) hoặc lõi IP. Dựa trên chức năng mô phỏng hoạt động truyền thông, lớp mạng cho phép tổng hợp thông tin cấu hình của toàn mạng làm cơ sở điều khiển hoạt động truyền thông thích ứng với sự thay đổi cấu hình. Do đó, ta có thể nhóm các công trình như: hệ điều hành cho hoạt động truyền thông mạng trên chip [51, 52], mạng quản lý trên chip (MNoC: *Monitor Network-on-Chip*) [74]; hoặc cơ chế điều chuyển dữ liệu Wormhole [17] đại diện cho khả năng quản lý cấu hình và điều khiển truyền thông.

2. ***Giao thức truyền thông linh hoạt***
   Theo nguyên lý truyền thông qua mạng, các hoạt động truyền thông sử dụng tập các giao thức truyền thông để thực hiện các yêu cầu truyền thông của ứng dụng và đồng thời quản lý hoạt động trao đổi thông tin giữa các thành phần cấu thành mạng. Sự linh hoạt của các giao thức truyền thông chính là ưu điểm của mạng, nên đây cũng là hướng nghiên cứu được chú trọng trong các giải pháp tái cấu hình mạng trên chip.
   Hướng đến khả năng thích ứng với sự thay đổi cấu hình mạng, các giải thuật định tuyến luôn được quan tâm nghiên cứu trong mạng trên chip tái

cấu hình. Một giải thuật định tuyến thích nghi cho phép lựa chọn đường định tuyến một cách linh hoạt, phù hợp với điều kiện mạng ngay cả trong khi hệ thống đang hoạt động. Flich và cộng sự đã phát triển nhóm giải pháp tái cấu hình dựa trên giải thuật định tuyến phân tán trong các công trình [18, 19, 45, 59]. Cùng định hướng sử dụng giải thuật định tuyến thích nghi cho hoạt động tái cấu hình nhưng áp dụng cho định tuyến tại nguồn bằng bảng tìm kiếm LUT, Mak và cộng sự đề xuất giải thuật *k-step look ahead* [41] và Miura đề xuất giải thuật Bắc-Nam trước (NSF: *North-South First routing*) cho dạng cấu trúc mạng lưới hai chiều 2D-torus [49].

Để hạn chế sự phức tạp của các giải thuật định tuyến thích nghi, các giải pháp cập nhật định tuyến cho phép thay đổi đường định tuyến tại một số vị trí đơn lẻ cũng được đề xuất như trong: giải thuật định tuyến kết hợp DyAD và định tuyến theo yêu cầu [27, 47]; hay giải thuật định tuyến đi vòng (SXY: *Surround XY routing*) [7, 6] và cập nhật định tuyến theo nguyên lý góc mở (RRAFT: *Reconfigurable Routing Algorithm for a Fault-Tolerant*) [73].

Cùng trong nhóm giải pháp phát triển giao thức truyền thông nhưng là sự kết hợp chức năng quản lý trạng thái mạng và giải thuật định tuyến tại lõi chức năng nhằm hướng đến mục tiêu nâng cao hiệu quả truyền thông cho toàn mạng, giải thuật định tuyến ATDOR do Manevich và cộng sự đề xuất trong hoạt động theo sự điều khiển tập trung của một lõi chức năng [42].

3. ***Kiến trúc tái cấu hình***

Các thành phần vật lý cấu thành kiến trúc mạng truyền thông có tác động sâu rộng đến toàn bộ khả năng đáp ứng các yêu cầu trao đổi thông tin của hệ thống. Cùng với đặc điểm thực thi trên chip, mọi hoạt động truyền thông qua mạng đều dựa trên các kiến trúc truyền thông phù hợp. Do vậy, giải pháp tái cấu hình của mạng trên chip chủ yếu tập trung chính vào các đề xuất mô hình kiến trúc cho phép thực hiện các hoạt động truyền thông linh hoạt. Một định hướng nghiên cứu rất được chú trọng của mạng trên chip tái cấu hình là khả năng linh hoạt của các tài nguyên truyền thông nhằm đáp ứng hiệu quả cho yêu cầu truyền thông của ứng dụng và tối ưu các chi phí thiết kế cũng như hoạt động của mạng. Các giải pháp tái cấu hình cho các tài nguyên truyền thông khá đa dạng, như khả năng thay đổi băng thông [33], tối ưu cho dung lượng bộ đệm trong các nút mạng [50] và chia sẻ tài nguyên bộ đệm bên trong bộ định tuyến [11]. Một trong những mục tiêu xây dựng kiến trúc tái cấu hình hướng đến khả năng thích ứng

cho sự thay đổi của các liên kết vật lý giữa các thành phần mạng, các giải pháp tái cấu hình cho cấu trúc liên kết đáp ứng cho mục tiêu đa ứng dụng hoặc chống sai hỏng như trong [61, 62] hoặc [71, 72].

2.3 Một số kiến trúc mạng trên chip tái cấu hình điển hình

Mạng trên chip là một giải pháp truyền thông của hệ thống trên chip nên cần hoàn thiện các chức năng bảo đảm đáp ứng tối ưu cho các yêu cầu trao đổi thông tin từ các lõi IP trên một vi mạch cụ thể. Do đó các giải pháp được đề xuất cho mạng trên chip tái cấu hình cần được cụ thể hóa bằng các kiến trúc thực thi bởi các mạch điện cụ thể trên vi mạch. Hơn nữa các giải pháp nghiên cứu các kiến trúc cho phép mạng trên chip có những khả năng linh hoạt thay đổi cấu hình để đáp ứng tốt hơn các nhiệm vụ truyền thông trong nhiều điều kiện khác nhau của hệ thống theo nhiều định hướng cụ thể. Do vậy, mục này trình bày một số các kết quả nghiên cứu kiến trúc mạng trên chip tái cấu hình điển hình. Với đối tượng nghiên cứu thiết kế là kiến trúc của các thành phần mạng trên chip, các giải pháp tái cấu hình hướng đến khả năng thích ứng với các thay đổi các thành phần vật lý cụ thể như: cấu trúc liên kết, bộ định tuyến hoặc các tài nguyên truyền thông đa dạng khác.

Các thành phần vật lý của mạng trên chip là nền tảng cơ bản và là phương tiện truyền thông cho mọi hoạt động trao đổi thông tin giữa các lõi IP; do đó mọi sự thay đổi thành phần truyền thông đều cần có các giao thức truyền thông hoặc giải pháp quản lý ở mức cao kết hợp. Và để bảo đảm khả năng truyền thông ổn định trong điều kiện thay đổi linh hoạt các thành phần của mạng, các kiến trúc của các thành phần vật lý cần được bổ sung các chức năng cụ thể nhằm tăng tính linh hoạt của hoạt động truyền thông.

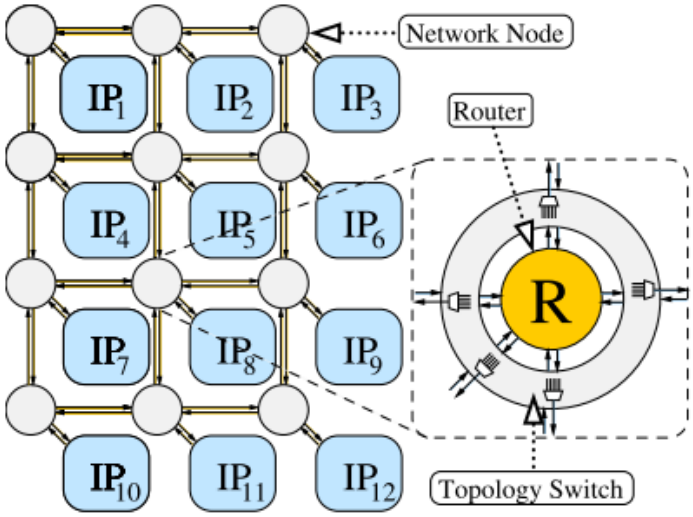
Với đặc điểm linh hoạt của các kỹ thuật truyền thông ở các lớp cao trong hoạt động đáp ứng yêu cầu trao đổi thông tin giữa các lõi IP của mạng trên chip dựa trên tài nguyên truyền thông của mỗi cấu hình mạng. Các giải pháp tái cấu hình truyền thông cho phép thực hiện các giao tác truyền thông một cách linh hoạt đáp ứng tối ưu cho các thay đổi của cấu hình mạng như đã được mô tả trong Hình 2.3. Do vậy, trong mục này luận án tập trung trình bày tổng quan các giải pháp hướng đến khả năng quản lý cấu hình và các giải pháp truyền thông linh hoạt cho phép thực hiện các mạng trên chip tái cấu hình.

2.3.1 Tái cấu hình cấu trúc liên kết

Ở mức thấp nhất của mô hình truyền thông qua mạng là các liên kết vật lý giữa các bộ định tuyến, các giải pháp tái cấu hình hướng đến khả năng tạo nên mạng trên chip có khả năng sắp xếp linh hoạt cấu trúc liên kết sẽ cho phép thích ứng tối đa cho mọi hoạt động của hệ thống. Tuy nhiên, giải pháp linh hoạt cấu trúc liên kết cần được xây dựng một cách đồng bộ từ kiến trúc bộ định tuyến và kết hợp với các giao thức truyền thông phù hợp với mọi sự thay đổi cấu hình và có thể cần cả những hoạt động quản lý cấu hình ở mức cao. Trong mục này trình bày tổng quan định hướng nghiên cứu này thông qua một số công trình điển hình đã công bố như sau:

Kiến trúc mạng tái cấu hình ReNoC, được nhóm nghiên cứu của Trường Đại học Kỹ thuật Đan Mạch phát triển cho phép linh hoạt sắp xếp các bộ định tuyến tạo nên các cấu trúc liên kết bất quy tắc từ cấu trúc dạng lưới hai chiều 2D-mesh [61, 62].

ReNoC sử dụng giải pháp bao bọc bên ngoài bộ định tuyến chuyển mạch gói là chuyển mạch kênh có khả năng thiết lập các liên kết vật lý một cách linh hoạt để sắp xếp các bộ định tuyến phù hợp với yêu cầu của hệ thống. Kiến trúc ReNoC với bộ định tuyến kết hợp hai dạng chuyển mạch (chuyển mạch kênh bao ngoài một chuyển mạch gói) được mô tả như trong Hình 2.4.



Hình 2.4: Kiến trúc mạng và bộ định tuyến của ReNoC [61].

Bằng cách bao ngoài bộ định tuyến chuyển mạch gói bởi bộ hợp kênh/phân kênh cho phép kết nối một cách linh hoạt các liên kết vật lý với các cổng vào/ra của bộ định tuyến cho phép thiết lập được nhiều dạng cấu trúc liên kết đáp ứng tối ưu cho nhiều ứng dụng khác nhau khi triển khai các hệ thống trên chip [61].

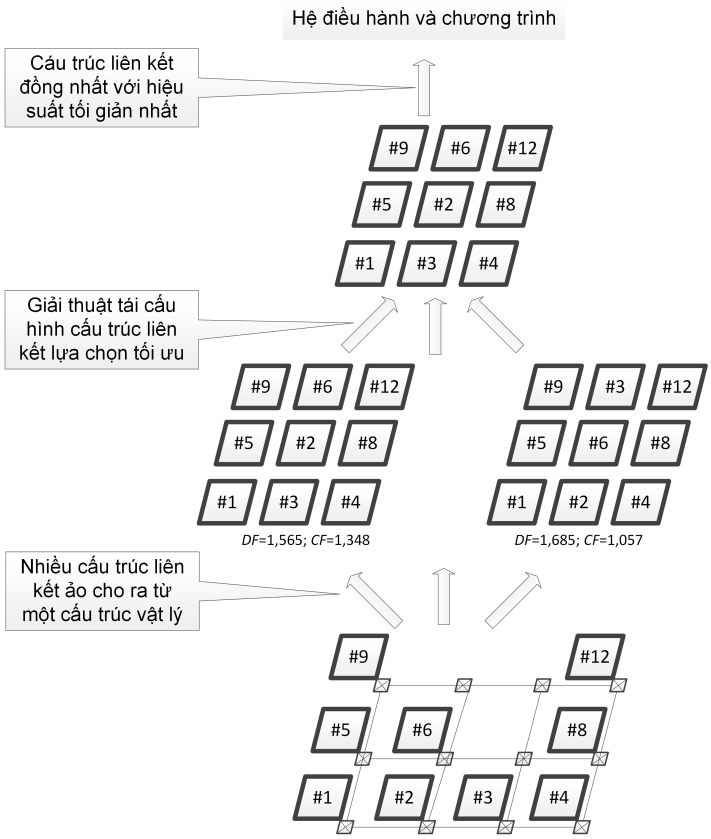
Để thiết lập được cấu trúc liên kết cho mạng trên chip tái cấu hình (ReNoC), nhóm tác giả hoàn thiện tập các giải pháp về thiết lập ban đầu cho hệ thống và khả năng định tuyến phụ thuộc các cấu trúc được thiết lập trên cơ sở bảng tìm kiếm LUT [62].

Cấu trúc liên kết ảo cho mục tiêu chống sai hỏng [71, 72]. Cùng hướng đến khả năng thích ứng với sự thay đổi của cấu trúc liên kết mạng (*nhưng lại tập trung vào khả năng chống sai hỏng trong quá trình sản xuất*) bằng sự linh hoạt của hệ điều hành mạng, Zhang và cộng sự đã phát triển giải pháp cấu trúc liên kết ảo (*Virtual topology*) dạng lưới hai chiều 2D-mesh. Trên nền tảng mạng sử dụng cấu trúc liên kết dạng lưới hai chiều 2D-mesh và hoạt động quản lý của hệ điều hành mạng cho phép tổng hợp trạng thái về cấu hình và truyền thông, nhóm tác giả đề xuất giải thuật tái sắp xếp các bộ định tuyến dựa trên hai thông số chính là: hệ số khoảng cách (DF: *Distance Factor*) và hệ số tắc nghẽn (CF: *Congestion Factor*) (Hình 2.5). Cấu trúc liên kết mạng sẽ được tái sắp xếp để tạo nên cấu trúc liên kết ảo dạng quy chuẩn 2D-mesh dựa trên khả năng cân bằng các giá trị DF và CF phù hợp với các ứng dụng và sau đó thiết lập thông tin định tuyến tại các bảng tra cứu thông tin định tuyến (LUT) ở mỗi nguồn phát. Phát triển giải pháp tái cấu trúc liên kết ảo, nhóm tác giả đề xuất thuật toán Bubble-Up để nâng cao hiệu quả truyền thông giữa các lõi chức năng khi kết hợp với các đặc tính truyền thông (C2C: *core-to-core*) trong hệ thống.

Mạng trên chip cấu hình CoNoChi (*Configurable Network-on-Chip*): Dựa vào khả năng linh hoạt của công nghệ FPGA, các giải pháp tái cấu hình mạng trên chip được nghiên cứu và phát triển theo mục tiêu tự thích ứng theo yêu cầu hoạt động của hệ thống. Điển hình là giải pháp mạng trên chip với khả năng thay đổi cấu hình một cách linh hoạt CoNoChi [55]. Giải pháp này cho phép hệ thống có thể tự thay đổi cấu hình mạng ngay cả khi đang hoạt động thông qua khả năng điều khiển hoạt động bổ sung/loại bỏ các bộ định tuyến và cập nhật bảng định tuyến thông qua bảng cấu hình động. Bằng cách đánh địa chỉ theo nhiều lớp và khả năng cập nhật bảng tìm kiếm LUT nhờ giải thuật Floy-Warshall sau khi cấu hình đã được thiết lập trên nguyên tắc tìm đường ngắn nhất giữa mỗi cặp nguồn đích.

2.3.2 Tái cấu hình kiến trúc bộ định tuyến

Hiện nay hệ thống trên chip được áp dụng nhiều cho các thiết bị thông minh cầm tay với nguồn năng lượng hạn chế và điều kiện hoạt động đa dạng nên cần những thiết kế hệ thống tối ưu để nâng cao hiệu quả hoạt động nhưng



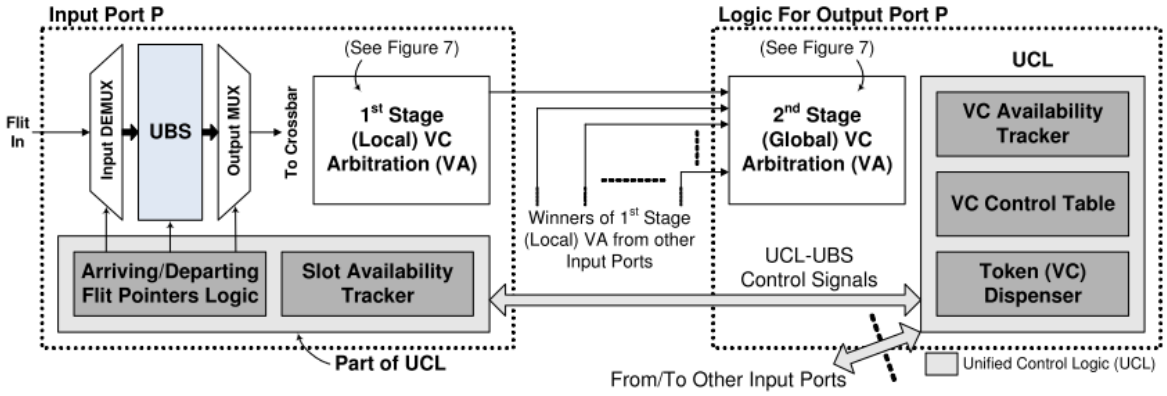
Hình 2.5: Hoạt động thiết lập cấu trúc liên kết ảo [71].

phải giảm nhỏ các chi phí thiết kế thực thi và tiêu hao năng lượng. Trong định hướng nghiên cứu mạng trên chip tái cấu hình đã có nhiều công trình hướng đến khả năng linh hoạt của truyền thông nhằm tối ưu tài nguyên và nâng cao hiệu quả truyền thông. Trong quá trình truyền thông các giao thức điều khiển truyền thông luôn phải dựa trên khả năng hoạt động của các tài nguyên truyền thông (khả năng đệm dữ liệu, phân xử yêu cầu truyền thông, băng thông ...) để đáp ứng các yêu cầu trao đổi dữ liệu từ các lõi chức năng. Do vậy, các giải pháp hướng trọng tâm đến hoạt động của bộ đệm bên trong các bộ định tuyến hoặc băng thông của các liên kết được gọi chung là các giải pháp quản lý tài nguyên truyền thông một cách linh hoạt.

Giải pháp bộ đệm thống nhất ViChar (*Virtual Channel Regulator*) do nhóm nghiên cứu tại trường Đại học bang Pennsylvania, Hoa Kỳ đề xuất [50].

Trong hoạt động truyền thông mạng trên chip, các bộ định tuyến sẽ nhận và đệm thông tin tại các khối nhớ ở cổng lối vào hoặc lối ra phụ thuộc vào cơ chế điều chuyển dữ liệu kết hợp với số lượng kênh ảo. Kiến trúc Vichar có khả năng cấp phát tài nguyên bộ đệm một cách linh hoạt cho tất cả các gói tin hướng đến bộ định tuyến và tối ưu chi phí thực hiện bộ đệm trong bộ định tuyến. Giải

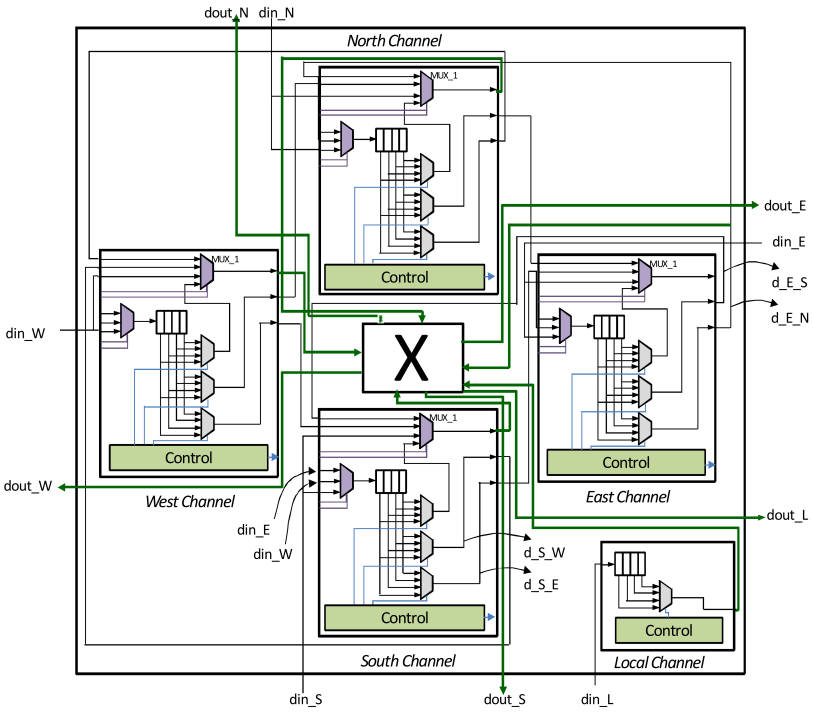
pháp cấu trúc bộ đệm thống nhất (UBS: *Unified Buffer Structure*) dùng cho cho tất cả các kênh ảo đầu vào của bộ định tuyến dưới sự điều khiển của khối logic thống nhất (UCL: *Unified Control Logic*) được mô tả trong Hình 2.6. Bằng việc kết hợp tài nguyên bộ đệm bên trong các bộ định tuyến, kiến trúc này cho phép tận dụng tối đa tài nguyên trong hoạt động truyền thông, đồng thời cho phép giảm chi phí thực thi (giảm gần 50% chi phí thực thi phần cứng) dẫn đến giảm công suất tiêu hao khoảng 30% so với bộ định tuyến cùng loại không sử dụng cấu trúc bộ đệm UBS như với ViChar.



Hình 2.6: Kiến trúc bộ đệm đa năng trong ViChar [50].

Trong hoạt động truyền thông mạng trên chip, để bảo đảm chất lượng dịch vụ truyền thông thì tất cả các lối vào đều cần được bố trí đầy đủ tài nguyên bộ đệm giống nhau. Nhưng trong quá trình hoạt động, thông tin hướng đến các bộ định tuyến rất đa dạng và có thể không cân bằng giữa các hướng khác nhau, lúc đó có thể xảy ra trường hợp một số cổng lối vào không thể cung ứng đủ bộ đệm cho các hoạt động truyền thông trong khi các khối cổng khác lại dư thừa tài nguyên. Concatto và cộng sự với mục tiêu tối ưu công suất trong hoạt động truyền thông mạng trên chip đã đề xuất giải pháp sử dụng bộ định tuyến tái cấu hình thông qua việc linh hoạt tài nguyên bộ đệm của bộ định tuyến (PRouter) [11]. Bộ định tuyến của mạng trên chip được bổ sung các khối ghép kênh ngay trước và sau khối bộ đệm để cho phép chia sẻ tài nguyên bộ đệm linh hoạt giữa các cổng lối vào nhằm mục tiêu giảm số lượng bộ đệm nhưng vẫn đáp ứng tối đa các hoạt động truyền thông của mạng (Hình 2.7).

Kết quả đạt được của nhóm kiến trúc linh hoạt tài nguyên bộ đệm cho thấy được khả năng tối ưu truyền thông của giải pháp tái cấu hình mạng trên chip hướng đến việc giảm thiểu chi phí thực thi và hoạt động của mạng nhưng vẫn bảo đảm đáp ứng các yêu cầu trao đổi thông tin của hệ thống. Tuy nhiên, cả hai giải pháp này lại tập trung hướng đến hoạt động truyền thông của mạng trên chip sử dụng các giao thức cần nhiều tài nguyên bộ đệm; nhưng điều này



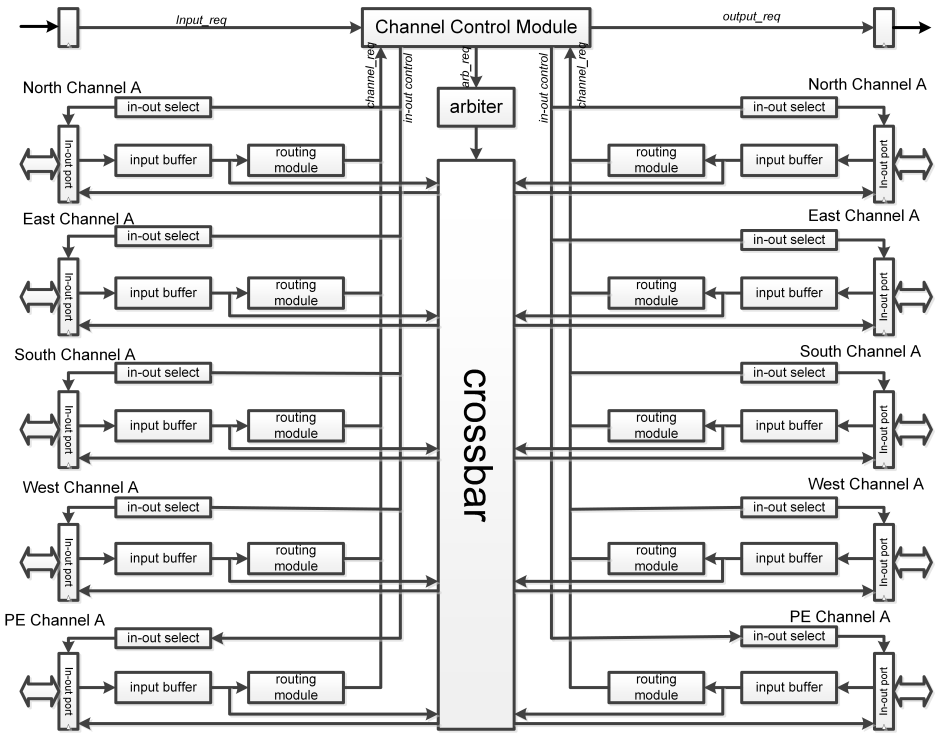
Hình 2.7: Kiến trúc bộ định tuyến chia sẻ tài nguyên bộ đệm [11].

chính là hạn chế lớn khi thực thi trên chip.

Kiến trúc mạng trên chip hai hướng BiNoC (*Bidirection Network-on-Chip*) do nhóm nghiên cứu kết hợp của Đại học quốc gia Đài Loan và Đại học Wisconsin, Hoa Kỳ đề xuất. Nghiên cứu này tập trung vào việc tăng sự linh hoạt trong sử dụng băng thông của liên kết mạng.

Trong hầu hết các hệ thống trên chip phức hợp, hoạt động trao đổi thông tin là ngẫu nhiên giữa các lõi chức năng. Với các thiết bị thông minh có tương tác, hoạt động truyền thông thường thay đổi động để phù hợp với nhiều chế độ hoạt động của hệ thống. Do vậy, giải pháp truyền thông cần có sự tương thích với hoạt động hệ thống thông qua khả năng cung cấp băng thông linh hoạt giữa các đối tượng. Đáp ứng yêu cầu này, Lan và cộng sự [33] đề xuất kiến trúc bộ định tuyến liên kết hai chiều có tên gọi là BiNoC như trong Hình 2.8.

Bằng cách bổ sung chức năng đệm hai chiều cho các liên kết vật lý giữa các thành phần của mạng trên chip, BiNoC cho phép thay đổi hướng truyền giữa các thành phần thông qua giao thức điều khiển hướng truyền được bắt tay giữa các bộ định tuyến. Ở chế độ mặc định, liên kết giữa các thành phần mạng là hai liên kết một chiều ngược hướng cho phép trao đổi thông tin song song hai chiều; nhưng khi có yêu cầu trao đổi đột biến một bộ định tuyến có thể yêu cầu thay đổi hướng truyền của liên kết với bộ định tuyến lân cận để đổi chiều liên kết cho phép tăng băng thông giữa hai bộ định tuyến.



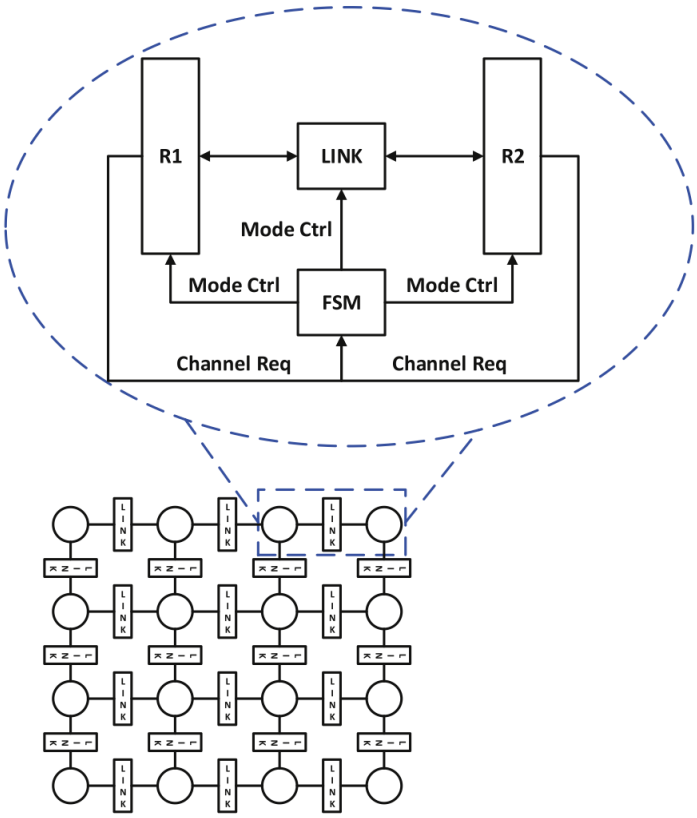
Hình 2.8: Kiến trúc bộ định tuyến liên kết hai chiều (BiNoC) [33].

Kiến trúc mạng trên chip sử dụng liên kết hai hướng BiLink (*Bi-directional link*) do nhóm nghiên cứu kết hợp giữa Đại học Khoa học và Công nghệ, Hồng Kông và Đại học Giao thông Thượng Hải, Trung Quốc phát triển với mục tiêu nhân đôi tốc độ và băng thông của các liên kết giữa các bộ định tuyến [75].

BiLink đưa ra giải pháp bổ sung hai đơn vị nhớ cho mỗi liên kết giữa các bộ định tuyến cũng như ở đầu ra của bộ định tuyến để cho phép các liên kết này có khả năng truyền dẫn hai chiều và hoạt động chỉ trong một pha của xung nhịp nhằm tăng gấp đôi tốc độ trao đổi dữ liệu. Kiến trúc mạng trên chip sử dụng kết nối hai chiều BiLink được mô tả như trong Hình 2.9. Hoạt động trao đổi thông tin giữa hai bộ định tuyến sử dụng chung liên kết hai chiều có một khối điều khiển tự tái cấu hình sử dụng nguyên tắc máy trạng thái hữu hạn để điều khiển hướng truyền thông tin.

2.4 Các vấn đề cần quan tâm khi xây dựng giải pháp truyền thông tái cấu hình

Qua phân tích các giải pháp mạng trên chip có khả năng tái cấu hình, ta có thể thấy rằng có 3 vấn đề cần quan tâm khi xây dựng một mạng trên chip có khả năng tái cấu hình đó là: ① Hạ tầng kiến trúc mạng cho phép thực thi



Hình 2.9: Liên kết nối giữa các bộ định tuyến trong dạng lưới hai chiều 2D-mesh [75].

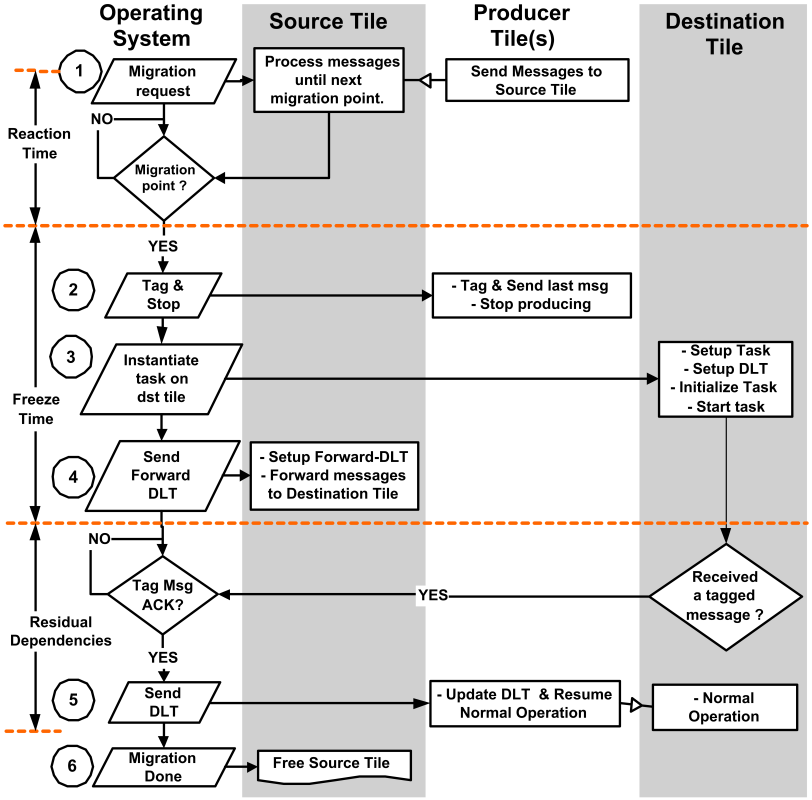
cập nhật cấu hình; ② Hoạt động quản lý cấu hình và điều khiển truyền thông; và ③ Giải thuật định tuyến cho mạng trên chip tái cấu hình. Mục trước đã đề cập khá chi tiết các kiến trúc mạng cho phép thực thi cập nhật tái cấu hình, do vậy, mục này tập trung vào hai vấn đề chính là hoạt động quản lý truyền thông và giải pháp định tuyến dùng cho các mạng trên chip tái cấu hình.

2.4.1 Hoạt động quản lý cấu hình và điều khiển truyền thông

Hoạt động nghiên cứu khả năng tái cấu hình mạng trên chip đặt ra vấn đề tiên quyết đó là làm thế nào để nhận biết được có sự thay đổi cấu hình rồi từ đó mới đưa ra những sự thích ứng phù hợp với các cấu hình thay đổi đó. Trong mô hình phân lớp của hoạt động truyền thông, lớp mạng với chức năng quản lý hoạt động của hạ tầng truyền thông cho phép thực hiện các giải pháp tổng hợp thông tin cấu hình và điều khiển toàn diện các hoạt động trao đổi thông tin qua mạng. Kết hợp với đặc điểm hoạt động của mạng trên chip như đã đề cập ở Chương 1, chức năng lớp mạng được thực hiện bởi giao tiếp mạng NI (*hoặc ngay tại lõi IP được kết nối trực tiếp với các bộ định tuyến*). Do vậy, các giải pháp tái cấu hình hướng đến mục tiêu quản lý cấu hình và điều khiển truyền thông chủ yếu được thực hiện bởi các lõi IP với chức năng hệ điều hành [52],

quản lý mạng [74], hoặc ở giao tiếp mạng [17].

Với hướng quản trị cấu hình đáp ứng linh hoạt cho các yêu cầu trao đổi thông tin của các ứng dụng, Nollet và cộng sự phát triển lõi IP có chức năng điều hành hoạt động của toàn mạng trên chip [52]. Giải pháp hệ điều hành mạng trên chip cung cấp khả năng điều khiển hoạt động truyền thông cả về lưu lượng thông tin và định tuyến thích nghi. Một ưu điểm nổi bật của hệ điều hành mạng là khả năng điều khiển một cách linh hoạt tài nguyên đáp ứng cho các giao tác truyền thông với giải pháp dịch chuyển tác vụ (Hình 2.10) [51].

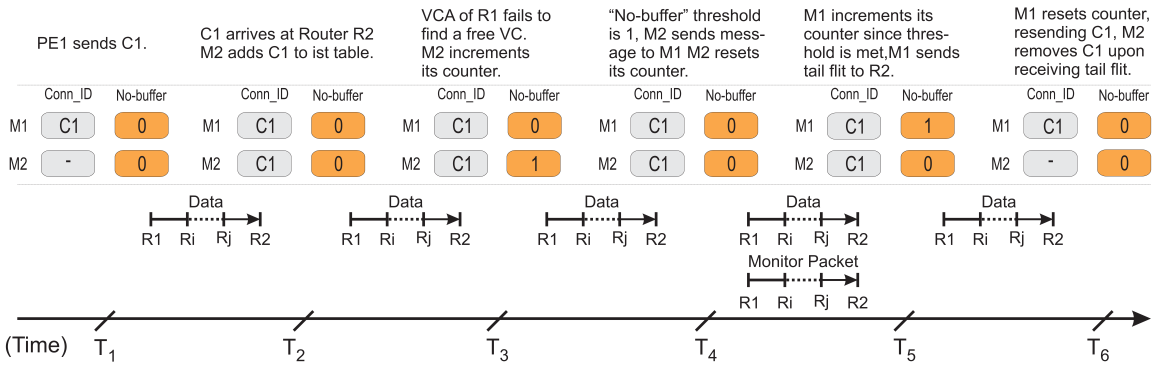


Hình 2.10: Giải pháp dịch chuyển tác vụ truyền thông [51].

Trong môi trường mạng các giao tác truyền thông chia sẻ tài nguyên truyền thông một cách linh hoạt dựa theo khả năng đáp ứng các yêu cầu từ ứng dụng; do vậy khi có sự thay đổi cấu hình sẽ ảnh hưởng sâu rộng đến các thành phần khác nhau của mạng. Ở mức tổng quát, mô hình mạng cho các nhu cầu truyền thông, các giải pháp quản lý mạng ở mức cao cho phép thu nhận trạng thái truyền thông và cân bằng tải tin cho toàn mạng thông qua khả năng điều khiển hoạt động truyền thông của các thành phần mạng một cách linh hoạt và hiệu quả.

Với đặc điểm của cơ chế truyền thông WH (xem Mục 1.3.2), giao dịch sẽ chiếm tài nguyên tại các bộ định tuyến trung gian và chỉ giải phóng khi gói tin tiến đến đích toàn vẹn. Tuy nhiên, trong nhiều trường hợp sai hỏng thông tin bị

chặn ở vị trí nào đó trên mạng và sẽ chiếm giữ toàn bộ tài nguyên của các bộ định tuyến trung gian làm ảnh hưởng lến đến các hoạt động truyền thông khác. Kiến trúc ROAdNoC [17] sử dụng nguyên tắc điều khiển của giao tiếp mạng và khả năng nhận biết sự cố ở bộ định tuyến để đề xuất giải pháp thu hồi tài nguyên của các giao dịch truyền thông bị chặn tại các bộ định tuyến như mô tả trong Hình 2.11. Khi một bộ định tuyến nhận được flit tiêu đề sẽ kích hoạt bộ đếm sự kiện thiết lập bên trong nó, bộ đếm sẽ tiến đến ngưỡng khi gói tin không thể chuyển tiếp hướng đến đích (*thời điểm T4*), giao tiếp mạng nhận được sự cố sẽ gửi gói tin giải phóng tài nguyên theo hướng ngược lại để hướng đến đích và yêu cầu một một flit đuôi để kết thúc giao dịch.



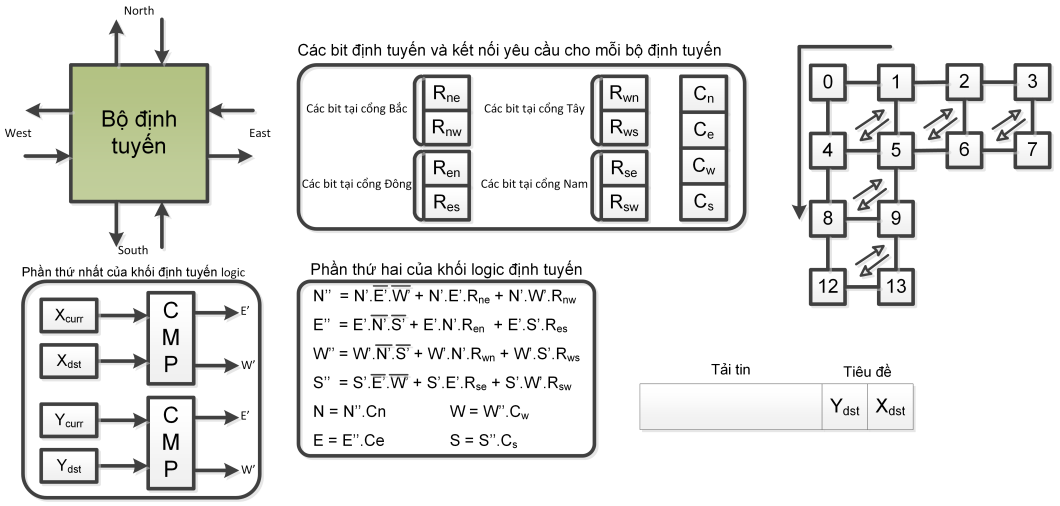
Hình 2.11: Giải pháp thu hồi tài nguyên của ROAdNoC [17].

2.4.2 Giải thuật định tuyến cho mạng trên chip tái cấu hình

Một khi kiến trúc truyền thông của mạng trên chip thay đổi (tái cấu hình) thì giải thuật định tuyến cũng cần thay đổi cho phù hợp với cấu hình mới của mạng. Tùy thuộc vào giải pháp tái cấu hình, người thiết kế cần đề xuất giải thuật định tuyến phù hợp, đồng bộ với giải pháp. Thường có hai xu thế: hoặc là xây dựng một thuật toán toàn năng phù hợp cho mọi thay đổi của kiến trúc mạng như giải thuật định tuyến thích nghi, hoặc là giải thuật định tuyến có khả năng tái cấu hình (thay đổi theo cấu hình mới của mạng). Với xu thế thứ nhất, giải thuật có thể áp dụng cho bất kỳ mô hình mạng nào và cấu hình nào. Nhược điểm là những thuật toán này rất phức tạp và khó triển khai hoặc chi phí quá lớn. Trong khi đó, xu thế thứ hai dễ được chấp nhận hơn vì dễ dàng thực thi với chi phí phù hợp. Hạn chế chủ yếu của những giải pháp này là chỉ áp dụng cho những kiến trúc cấu hình cụ thể.

Hướng đến trọng tâm là khả năng chống sai hỏng trong sản xuất vi mạch, nhóm tác giả từ Đại học Valencia – Tây Ban Nha gồm Jose Flich và nhiều cộng sự đã nghiên cứu và phát triển nhóm giải pháp định tuyến linh hoạt, thích ứng

với nhiều cấu trúc liên kết tùy biến từ cấu trúc dạng lưới 2D-mesh. Trên đặc điểm cấu trúc 2D-mesh, các sai hỏng mạng tạo nên các cấu trúc bất quy tắc nhưng vẫn có thể chia thành các đoạn nhỏ dạng khối chữ nhật [45]. Nhóm nghiên cứu đã kết hợp khả năng phân đoạn mạng (SR: *Segment-based Routing*) để đề xuất giải thuật định tuyến trên cơ sở phân miền (RR: *Region-based Routing*) [19]. Nguyên lý định tuyến của RR là ưu tiên chuyển tin bên trong mỗi miền, và sau đó hướng đến miền ngoài thông qua các kết nối khả dĩ dựa trên thông tin trạng thái được lưu ở mỗi bộ định tuyến và đích đến. Hoàn thiện hơn giải pháp định tuyến Flich và cộng sự phát triển giải thuật định tuyến phân tán trên cơ sở mạch logic (LBDR: *Logic-based Distributed Routing*) [18]. Giải pháp định tuyến LBDR là sự kết hợp giải thuật định tuyến RR với khả năng xử lý định tuyến dựa trên mạch logic và trạng thái liên kết giữa các bộ định tuyến như được mô tả trong Hình 2.12.

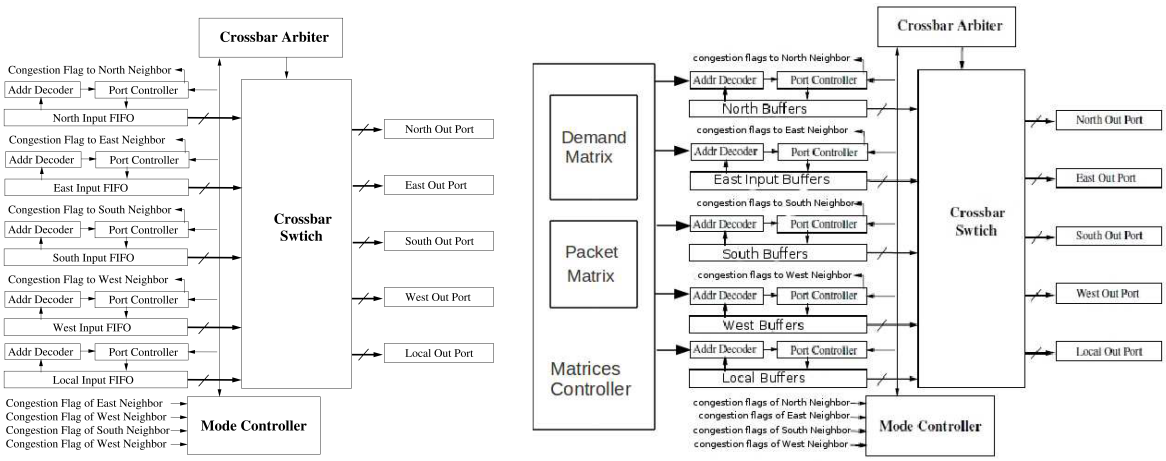


Hình 2.12: Giải pháp định tuyến phân tán trên cơ sở logic (LBDR) [19].

Với khả năng thích ứng tốt các thay đổi cấu hình mạng và sự linh hoạt của giải thuật định tuyến phân tán, nhóm tác giả phát triển bổ sung các tính năng mới đáp ứng cho các hệ thống phức hợp đa xử lý [20, 46, 58, 59].

Với những điều kiện ràng buộc khi thực thi trên vi mạch, mạng trên chip cần những giải thuật định tuyến đơn giản, hiệu quả và đang được phát triển đa dạng. Tuy nhiên trong định hướng tái cấu hình động với sự thay đổi linh hoạt cấu hình mạng thì không thể sử dụng được các giải thuật định tuyến tĩnh mà cần có sự phát triển bổ sung phù hợp. Đáp ứng yêu cầu truyền thông linh hoạt của mạng trên chip để cân bằng hoạt động truyền thông chống tắc nghẽn, giải pháp định tuyến thông minh DyAD [27] là sự kết hợp của định tuyến tĩnh XY và định tuyến thích nghi chẵn-lẻ (OE) [10]. Hu và cộng sự giới thiệu kiến trúc bộ định tuyến có khả năng thay đổi chế độ định tuyến dựa trên cờ trạng

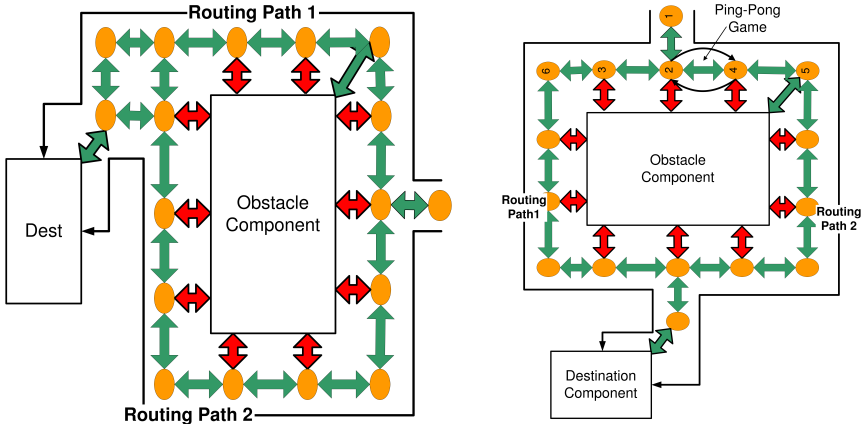
thái tắc nghẽn từ các bộ định tuyến lân cận và khối điều khiển như được mô tả trong Hình 2.13a [27]. Sau đó, Meka và cộng sự bổ sung khả năng tính toán yêu cầu truyền thông từ các lõi IP và trạng thái truyền thông bằng các ma trận nhớ trạng thái như trong Hình 2.13b [47]. Bằng giải pháp bổ sung các ma trận trạng thái và giao dịch thu nhận yêu cầu để tổng hợp trạng thái, giải pháp định tuyến theo yêu cầu hướng đến khả năng cân bằng tải tin trong toàn mạng.



(a) Kiến trúc bộ định tuyến DyAD [27] (b) Giải pháp định tuyến theo yêu cầu [47]

Hình 2.13: Nguyên lý kiến trúc bộ định tuyến kết hợp DyAD.

Cũng như đối với các hệ thống trên chip tái cấu hình, nhiều công trình nghiên cứu tái cấu hình mạng trên chip sử dụng công nghệ FPGA là nền tảng thực hiện các kiến trúc đề xuất. Vì đặc điểm cấu trúc của FPGA thường bố trí các tài nguyên có dạng lưới hai chiều và phân bố đều nên hỗ trợ tối ưu cho các thiết kế hệ thống sử dụng mạng trên chip có dạng lưới hai chiều 2D-mesh. Tuy nhiên, trong nhiều trường hợp thiết kế hệ thống trên chip bằng FPGA có những lõi IP chiếm kích thước lớn gây khó khăn trong việc bố trí các thành phần, nhất là sắp xếp cấu trúc của mạng trên chip. Bobda và cộng sự đã đề xuất giải thuật định tuyến đi vòng (S-XY: *Surround XY routing*) làm cơ sở cho mạng trên chip tái cấu hình động (DyNoC: *Dynamic Network-on-Chip*) [6, 7]. Giải thuật định tuyến S-XY là sự cải tiến của giải thuật định tuyến tĩnh XY khi bổ sung khả năng đi vòng qua các lõi chức năng rất lớn trong hai trường hợp như được mô tả trong Hình 2.14. Để hạn chế các trường hợp đường đi vòng quá dài, giải pháp DyNoC đề xuất thiết lập hướng truyền ưu tiên bằng cờ trạng thái cho hướng truyền tối ưu tại bộ định tuyến cập nhật đường định tuyến.



(a) Định tuyến vòng theo phương ngang (b) Định tuyến vòng theo phương dọc

Hình 2.14: Giải thuật định tuyến vòng trong DyNoC [7].

2.5 Tổng kết chương

Nhằm mục tiêu tổng quan hoạt động nghiên cứu các giải pháp thiết kế hệ thống trên chip trong xu thế tái cấu hình để cho một cách nhìn tổng thể những hướng nghiên cứu đang được quan tâm phát triển hiện tại. Các nghiên cứu giải pháp tái cấu hình cho phép nâng cao hiệu quả của thiết kế khi mang lại nhiều tính năng ưu việt bằng việc nâng cao khả năng ứng dụng cũng như hiệu quả trong hoạt động. Luận án này cũng đi sâu tìm hiểu các động lực và thách thức của vấn đề truyền thông trên chip cho các hệ thống phức hợp trong xu thế thiết kế lấy truyền thông làm trung tâm. Với xu thế thiết kế truyền thông làm trung tâm, mỗi thiết kế mạng trên chip tái cấu hình là sự tổng hợp nhiều lõi IP trên nền một hạ tầng truyền thông đủ mạnh và linh hoạt sẽ mang lại nhiều lợi ích khi giảm nhỏ chi phí nghiên cứu và thực thi hệ thống. Để có cái nhìn đầy đủ về mạng trên chip tái cấu hình Mục 2.2 đã tập trung làm rõ định hướng nghiên cứu của các giải pháp truyền thông mạng trên chip tái cấu hình và các cơ sở phân loại cho phép chúng ta có thể sắp xếp những giải pháp tái cấu hình trên chip trong một bối cảnh toàn diện và đầy đủ hơn. Trên cơ sở nghiên cứu những định hướng tái cấu hình mạng trên chip, chương này đã tổng quan một số công trình kiến trúc truyền thông mạng trên chip tái cấu hình để minh họa rõ nét hơn các mục tiêu phát triển của xu thế này. Đồng thời với việc đánh giá các kiến trúc mạng trên chip tái cấu hình, trong chương này cũng đã nêu bật được những vấn đề cốt lõi của những giải pháp tái cấu hình mạng trên chip thông qua việc đánh giá lại những kết quả của các giải pháp tái cấu hình trong hoạt động truyền thông linh hoạt để thích ứng tối đa khi hệ thống có sự thay đổi cấu hình. Và

để làm rõ hơn những thành quả đã đạt được của những công trình nghiên cứu về mạng trên chip tái cấu hình, những đặc điểm và thông số cơ bản của những công trình điển hình được tổng hợp trong Bảng 2.1. Nội dung tổng quan xu thế tái cấu hình và mạng trên chip tái cấu hình đã được tác giả và nhóm nghiên cứu công bố tại công trình [J2] để làm rõ hơn định hướng phát triển của các thiết kế tái cấu hình.

Thông qua việc tìm hiểu xu thế tái cấu hình của các thiết hệ thống trên chip để hướng đến những giải pháp mạng trên chip tái cấu hình, mà trọng tâm là hoạt động truyền thông linh hoạt có khả năng tự thích ứng với sự thay đổi cấu hình của hệ thống cho thấy đây là một định hướng phát triển rất đáng được quan tâm nghiên cứu. Trên cơ sở cung cấp khả năng truyền thông linh hoạt của mô hình mạng, các giải pháp tái cấu hình mạng trên chip đã tận dụng tối đa các khả năng linh hoạt của truyền thông với nhiều chức năng của các lớp khác nhau của mô hình hệ thống mở. Tuy nhiên, với đặc điểm thực thi trên chip sẽ ràng buộc các giải pháp tái cấu hình sử dụng những chức năng của các lớp cao với các hoạt động phức tạp sẽ ảnh hưởng lớn đến chi phí thiết kế và thực thi. Do vậy, giải pháp tái cấu hình mạng trên chip cần được xem xét một cách toàn diện hơn và tối ưu hơn trong quá trình triển khai trong khả năng cân bằng giữa hiệu năng hoạt động tái cấu hình và chi phí nghiên cứu và thực thi giải pháp đó trên chip. Một giải pháp tái cấu hình mạng trên chip tối ưu hơn bằng khả năng kết hợp sự linh hoạt của hoạt động định tuyến và hiệu quả hoạt động ổn định của kiến trúc bộ định tuyến tái cấu hình làm một hướng nghiên cứu rất thú vị và cần được quan tâm hơn.

Bảng 2.1: Bảng tóm tắt các công trình mạng trên chip tái cấu hình.

Công trình	**Đối tượng và mục tiêu**	**Cấu trúc liên kết**	**Hoạt động định tuyến**	**Kiến trúc mới**	**Kết quả thực thi cho bộ định tuyến**
ReNoC [61, 62]	Khả năng tự sắp xếp lại cấu trúc liên kết mạng để thích ứng với nhiều ứng dụng	Tùy biến bất quy tắc từ 2D-mesh	Giải thuật định tuyến cụ thể theo cấu trúc liên kết mạng	Chuyển mạch kênh bao ngoài bộ định tuyến	Công nghệ $90nm$: $580000\mu m^2$
Cấu trúc liên kết ảo [72]	Khả tái sắp xếp cấu trúc liên kết Thích ứng với sai hỏng trong sản xuất	Dạng lưới 2D-mesh	Giải thuật định tuyến tại nguồn bằng LUT được cập nhật bằng hệ điều hành		N/A
CoNoChi [55]	Khả năng thay đổi linh hoạt cấu hình mạng ngay khi đang hoạt động	Bất quy tắc	Cập nhật thông tin của bảng định tuyến LUT		Virtex-II Pro: $493 Slice$
ViChar [50]	Nâng cao hiệu quả sử dụng bộ đệm	Dạng lưới hai chiều	Định tuyến phân tán dựa theo đích đến	Bộ đệm thống nhất (UBS)	Công nghệ TSMC $90nm$: $101899\mu m^2$
PRouter [11]	Chia sẻ bộ đệm trong bộ định tuyến để giảm công suất tiêu hao	Dạng lưới hai chiều	Định tuyến tĩnh XY	Bổ sung ghép kênh ở cổng lối vào	Công nghệ CMOS $180nm$: $337471\mu m^2$, công suất $9{,}34mW$

BiNoC [33]	Linh hoạt băng thông giữa các bộ định tuyến để thích ứng với ứng dụng	Dạng lưới hai chiều	Định tuyến tĩnh XY	Kiến trúc bộ định tuyến liên kết hai chiều	Công nghệ TSMC $90nm$: $48666 gates$
BiLink [75]	Nhân đôi băng thông liên kết	Dạng lưới hai chiều	Định tuyến tĩnh XY	Kiến trúc bộ định tuyến và liên kết vật lý	Công nghệ TSCM $65nm$: $219470\mu m^2$
OSfNoC [52, 51]	Hoạt động quản lý cấu hình mạng và điều khiển hoạt động truyền thông thích ứng với cấu hình hiện tại	Dạng lưới hai chiều	Định tuyến thích nghi trên cơ sở hệ điều hành mạng	Lõi chức năng điều hành	N/A
ROAd NoC [17]	Quản lý hoạt động truyền thông qua mạng	Dạng lưới hai chiều	Định tuyến thích nghi dùng LUT	Khối điều khiển ở bộ định tuyến và NI	N/A
MNoC [74]	Quản lý cấu hình qua các cảm biến tích hợp	Dạng lưới hai chiều 2D-mesh	Sử dụng quản lý cấu hình để điều khiển bảng định tuyến	Kiến trúc lõi chấp hành MEP	N/A
AT-DOR [42]	Linh hoạt giải thuật định tuyến để cân bằng tải mạng	Dạng lưới hai chiều	Giải thuật ATDOR là sự kết hợp giữa định tuyến XY và YX	Khối điều khiển trung tâm	N/A

LBDR [20]	Sử dụng giải thuật định tuyến phân tán để thích ứng với sự thay đổi cấu hình	Tùy biến từ 2D-mesh	Giải thuật định tuyến phân tán trên cơ sở mạch logic	Kiến trúc cho định tuyến phân tán - LBDR	Công nghệ STMicroelectronics $65nm$: $20335\mu m^2$
KLSA [41]	Nâng cao hiệu quả truyền thông bằng giải pháp tối ưu đường định tuyến	Dạng lưới 2D-mesh	Định tuyến thích nghi trên cơ sở cập nhật LUT	Khối tính toán tối ưu đường truyền	Công nghệ FPGA VIRTEX-4 XC4VLX40: $384 Slice$
DyAD [27, 47]	Linh hoạt khả năng định tuyến để cân bằng tải mạng chống tắc nghẽn	Dạng lưới 2D-mesh	Định tuyến kết hợp	Khối điều khiển chế độ định tuyến	Công nghệ $160nm$: $25971 gates$
DyNoC [6, 7]	Linh hoạt cho giải thuật định tuyến tĩnh XY thích ứng với các lõi chức năng rất lớn	Định tuyến đi vòng SYX	Tùy biến 2D-mesh	N/A	Công nghệ FPGA Virtex-II 1000: 21%
RRAFT [73]	Thu nhận thông tin cấu hình để linh hoạt đường định tuyến cho giải thuật XY cho khả năng chống sai hỏng	Tùy biến 2D-mesh	Cập nhật định tuyến đi vòng		Công nghệ CMOS $90nm$: $680000\mu m^2$

Chương 3

Giải pháp tái cấu hình cho mạng trên chip

Hệ thống trên chip ngày càng phát triển đa dạng, tích hợp nhiều chức năng hơn và hoạt động cũng phức tạp hơn nhằm đáp ứng yêu cầu của ứng dụng. Với mục tiêu tiết giảm thời gian và chi phí nghiên cứu cũng như gia tăng đặc tính tối ưu cho các hệ thống thì xu thế thiết kế tái cấu hình đã và đang được tập trung nghiên cứu để cho ra nhiều sản phẩm nổi bật. Hệ thống tái cấu hình thích hợp cho những thiết kế hướng đến dải ứng dụng rộng, có thay đổi trong hoạt động thực tiễn. Nhờ đặc điểm tự thích ứng với nhiều điều kiện hoạt động cũng như khả năng tái sắp xếp các thành phần để hệ thống hoạt động tối ưu, hệ thống tái cấu hình cho phép nâng cao hiệu quả thiết kế khi tái sử dụng thiết kế trong nhiều ứng dụng hoặc chống lỗi trong sản xuất vi mạch.

Khi hệ thống được tích hợp nhiều lõi IP hơn và hoạt động cũng phức tạp hơn thì yêu cầu truyền thông bên trong hệ thống cũng cần được phát triển. Do đó, mạng trên chip được đề xuất và phát triển với nhiều đặc điểm nổi trội và cung cấp môi trường truyền thông hiệu quả trong các hệ thống tích hợp. Nhiều nghiên cứu gần đây đã chứng tỏ việc sử dụng mô hình mạng để giải quyết bài toán truyền thông trên chip của các hệ thống phức hợp là tối ưu và triệt để, khi mạng trên chip cung cấp khả năng truyền tốc độ cao, hỗ trợ xử lý song song thời gian thực. Hơn nữa, các nghiên cứu thiết kế mạng trên chip cũng góp phần hình thành xu thế thiết kế lấy truyền thông làm trung tâm khi dựa trên khả năng linh hoạt của mô hình mạng trong việc cung cấp các dịch vụ truyền thông cho đa dạng các dịch vụ. Do vậy, định hướng nghiên cứu thiết kế cho truyền thông trên chip để hướng đến một giải pháp toàn diện áp dụng cho các hệ thống phức hợp cần kết hợp cả hai xu thế trên để hướng đến các giải pháp tái cấu hình mạng trên chip.

Các giải pháp tái cấu hình cho mạng trên chip là rất đa dạng, hướng đến

các mục tiêu khác nhau (chống sai hỏng, cân bằng tải truyền thông, ...) và cũng tác động đến hoạt động truyền thông với nhiều góc độ khác nhau nhằm giải quyết một cách tối ưu các vấn đề truyền thông của hệ thống trên chip. Với mục tiêu xây dựng mô hình mạng trên chip có thể tùy biến theo yêu cầu cụ thể của ứng dụng, đáp ứng các hỏng hóc trong quá trình sản xuất hay lỗi do quá trình hoạt động, chương này sẽ trình bày giải pháp cho phép tái cấu hình mạng trên chip. Với giải pháp này, mạng trên chip có thể được tái cấu hình để hoạt động theo chủ ý của người thiết kế tùy vào bối cảnh triển khai ứng dụng hoặc duy trì hoạt động khi có hỏng hóc, lỗi xảy ra. Do vậy, nội dung chính của chương được chia thành các vấn đề cơ bản sau: cơ sở thực hiện giải pháp, giải thuật cập nhật thông tin định tuyến và triển khai giải pháp thông qua việc xây dựng các kiến trúc chi tiết mô tả chức năng hệ thống đến kiến trúc khả thi cho phép thực thi trên vi mạch.

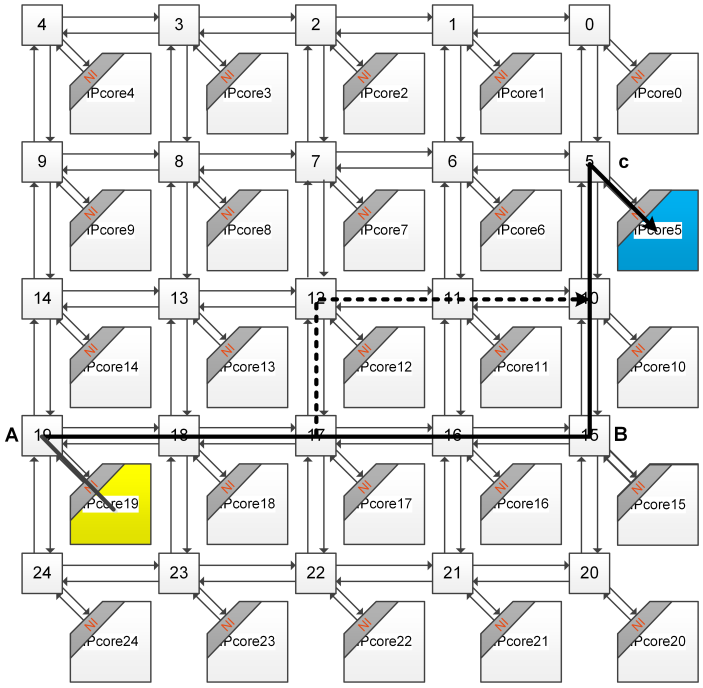
3.1 Cơ sở thực hiện giải pháp tái cấu hình

Xu hướng nghiên cứu mạng trên chip tập trung chính ở cấu trúc liên kết dạng lưới hai chiều 2D-mesh và áp dụng thuật toán định tuyến tĩnh. Ưu điểm của thuật toán định tuyến tĩnh là sự đơn giản trong thực thi (như đã nêu ở Mục 1.4.2), tuy nhiên, nhược điểm của các thuật toán này là hệ thống không thể tự động thay đổi đường truyền trong lúc hoạt động khi có lỗi xảy ra. Vậy nên, giải pháp tái cấu hình cần kết hợp được ưu điểm của giải pháp định tuyến tĩnh và khả năng thay đổi đường định tuyến một cách linh hoạt theo yêu cầu của hệ thống. Điều này sẽ cho phép tận dụng tính ưu việt của giải pháp định tuyến tĩnh ở tính đơn giản và khả năng cân bằng tải mạng, đồng thời bổ sung khả năng tự thích ứng với sự thay đổi cấu hình. Trước khi trình bày chi tiết giải pháp đề xuất, trong mục này sẽ trình bày một số định nghĩa và khái niệm cơ bản sau đây.

3.1.1 Một số định nghĩa

Xét hoạt động truyền thông của mạng trên chip có cấu trúc dạng lưới hai chiều 2D-mesh với giải thuật định tuyến tĩnh XY (xem Mục 1.4.2); đường định tuyến cho giao tác truyền thông giữa cặp ***nguồn-đích*** bất kỳ là sự kết hợp của các đoạn thẳng định tuyến và góc định tuyến. Hình 3.1 mô tả mạng trên chip có kích thước 5×5 gồm 25 bộ định tuyến với đường định tuyến được thiết lập (*kẻ đậm*) từ nguồn là lõi chức năng IPcore19 đến đích là lõi chức năng IPcore5. Để

thống nhất trong trình bày các thuật ngữ và giải pháp thực hiện tái cấu hình mạng trên chip, mục này sẽ đưa ra một số định nghĩa sau:



Hình 3.1: Mạng trên chip dạng lưới hai chiều kích thước 5×5 với thuật toán định tuyến XY.

Định nghĩa 1: ***Đường định tuyến*** là tập hợp tất cả k nút mạng trung gian mà gói tin cần đi qua cho phép giao tác truyền tin chuyển dữ liệu từ nguồn tin đến đích một cách tin cậy và hiệu quả dựa theo một giải thuật định tuyến cụ thể phù hợp với hoạt động truyền thông của toàn hệ thống. Trong trường hợp giải thuật định tuyến tĩnh, đường định tuyến giữa mỗi cặp nguồn-đích là cố định.

Định nghĩa 2: ***Đoạn thẳng định tuyến*** là một phần của đường định tuyến mà trên đó thông tin được truyền theo một hướng nhất định (Bắc, Đông, Nam hoặc Tây). Trên đoạn thẳng định tuyến thông tin đến và đi qua mỗi bộ định tuyến đều theo một phương nhất định (phương ngang hoặc phương dọc). Như trong Hình 3.1, có hai đoạn thẳng định tuyến là đoạn từ bộ định tuyến 19 đến bộ định tuyến 15 (đoạn AB) và đoạn từ bộ định tuyến 15 đến bộ định tuyến 5 (đoạn BC).

Định nghĩa 3: ***Góc định tuyến*** là điểm chuyển hướng trên đường định tuyến của gói tin; thông tin từ phương ngang (Đông-Tây) sẽ chuyển theo phương dọc (Bắc-Nam) và ngược lại. Trong trường hợp sử dụng giải thuật định tuyến tĩnh XY thì góc định tuyến là những điểm chuyển hướng các phương ngang sang phương dọc, hoặc tại các bộ định tuyến kết nối với đích đến của thông tin.

Góc định tuyến trong trường hợp (Hình 3.1) nằm tại vị trí bộ định tuyến 15, tương ứng với vị trí đường định tuyến chuyển hướng từ phương ngang sang phương dọc.

Trong định hướng nghiên cứu giải pháp tái cấu hình cho mạng trên chip, các trạng thái hoạt động của các bộ định tuyến cần được quan tâm và xem xét để đưa ra quyết định thay đổi cấu hình. Do vậy, trong phần sau đây sẽ định nghĩa các trạng thái hoạt động của bộ định tuyến của mạng trên chip tái cấu hình:

Định nghĩa 4: ***Bộ định tuyến bị cấm*** là bộ định tuyến không thể tham gia hoạt động truyền thông do có sai hỏng trong quá trình sản xuất hoặc do yêu cầu hoạt động của hệ thống và được thiết lập để rời khỏi mạng.

Định nghĩa 5: ***Bộ định tuyến tái cấu hình*** là bộ định tuyến mà ở đó hoạt động tái cấu hình được triển khai để thực hiện chức năng cập nhật thông tin định tuyến cho giao tác truyền thông nhằm tránh bộ định tuyến bị cấm.

Theo nguyên lý định tuyến tại nguồn kết hợp với giải thuật định tuyến tĩnh, đường định tuyến của mỗi giao tác truyền thông được tạo lập tại nguồn tin bằng cách thiết lập thông tin định tuyến phù hợp tại trường định tuyến (PTT) như trong Hình 3.3. Do vậy, hoạt động tái cấu hình cần thay đổi giá trị của trường PTT của *flit* tiêu đề để thay đổi đường định tuyến mới thích ứng với sự thay đổi của cấu hình mạng.

Khái niệm: Hoạt động cập nhật định tuyến là quá trình xử lý thay đổi thông tin định tuyến trong trường PTT của flit tiêu đề của gói tin để thay đổi nội dung của trường này cho phép giao tác truyền thông hướng đến đích theo một đường định tuyến mới phù hợp với cấu hình mạng hiện thời.

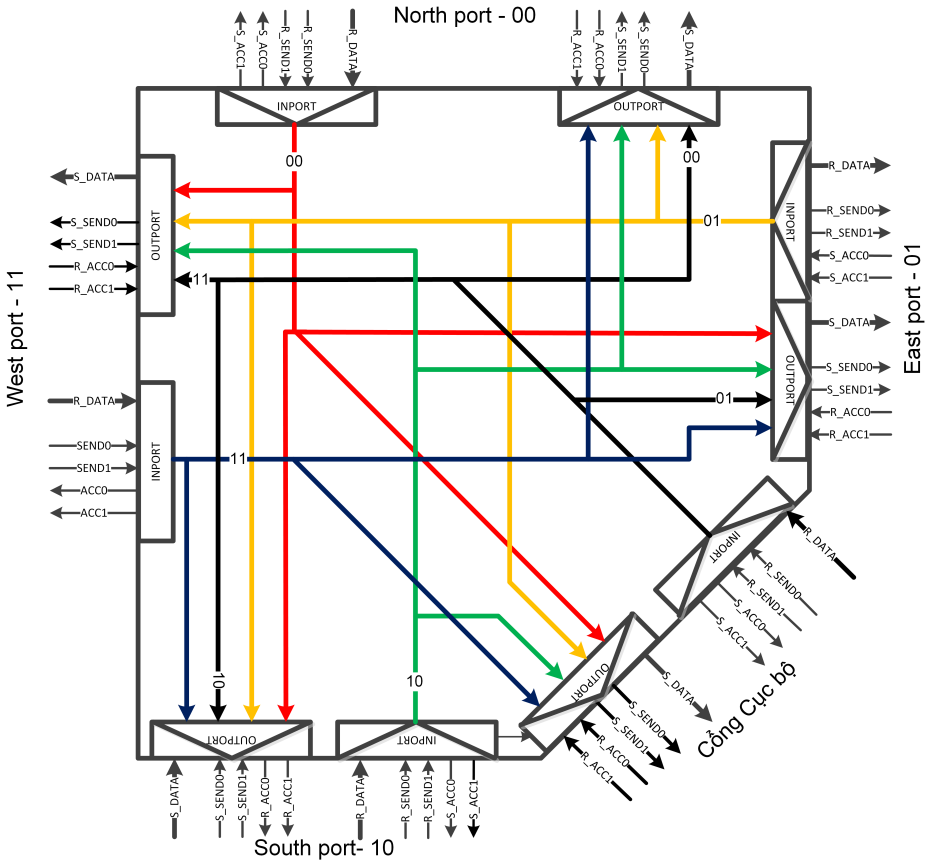
3.1.2 Cơ sở giải pháp tái cấu hình

Trên cơ sở nghiên cứu tổng quan hoạt động truyền thông mạng trên chip đã được trình bày ở Chương 1, giải pháp tái cấu hình được đề xuất hướng đến khả năng áp dụng cho kiến trúc truyền thông mạng trên chip đang được quan tâm nghiên cứu nhiều trong những năm gần đây với những đặc điểm truyền thông như sau:

- Mạng trên chip sử dụng kiến trúc truyền thông mạng có cấu trúc liên kết hai chiều dạng lưới (2D-mesh, 2D-torus).
- Hoạt động truyền thông mạng trên chip sử dụng kỹ thuật chuyển mạch gói kết hợp với cơ chế điều chuyển dữ liệu *wormhole* và cơ chế điều khiển

luồng *credit-based* (xem Mục 1.3).

- Giải thuật định tuyến sử dụng thuộc nhóm định tuyến tĩnh ưu tiên theo hướng (DOR) với cơ chế định tuyến tại nguồn (xem Mục 1.4). Cụ thể, trong luận án này sử dụng thuật toán định tuyến tĩnh XY tại nguồn để khảo sát. Theo đó, bộ định tuyến căn bản cho mạng trên chip dạng lưới hai chiều gồm năm cổng vào/ra được định danh bằng 2 *bit* tương ứng như trong Hình 3.2. Mỗi cổng lối vào được nối với 4 cổng lối ra còn lại, thông tin từ cổng lối vào sẽ hướng đến cổng lối ra dựa vào 2 *bit* thông tin định tuyến đã được thiết lập tại nguồn, trùng với định danh của cổng lối ra. Muốn gửi gói dữ liệu đến cổng cục bộ, thông tin định tuyến phải được lập bằng chính định danh của cổng lối vào.

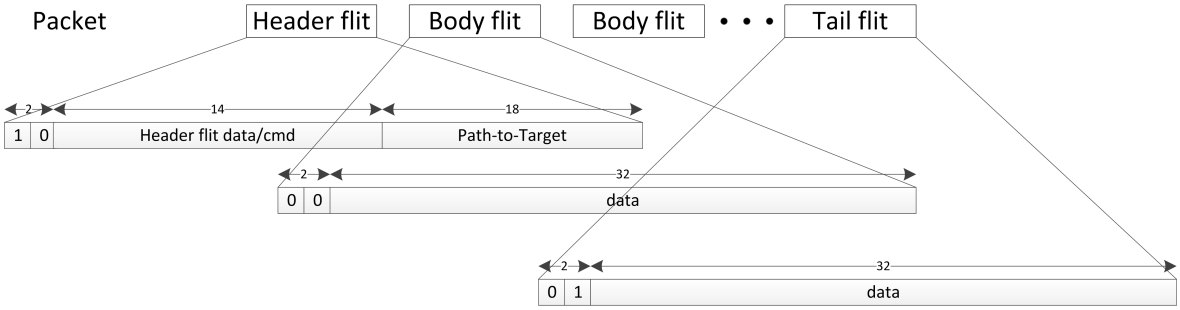


Hình 3.2: Bộ định tuyến căn bản cho mạng trên chip 2D-mesh.

Dựa theo ví dụ mô tả trong Hình 3.1, giả thiết rằng bộ định tuyến 16 rời khỏi mạng sẽ là bộ định tuyến bị cấm và giao tác truyền thông sẽ bị chặn tại bộ định tuyến 17 do cổng lối ra hướng **Đông** bị chặn (liên kết với cổng lối vào của bộ định tuyến bị cấm). Để thích ứng với sự thay đổi cấu hình mạng do bộ định tuyến 16 bị cấm, bộ định tuyến 17 cần xử lý lại thông tin định tuyến. Với giao tác truyền thông nói trên, tại nguồn tin, trường

định tuyến có giá trị nhị phân sẽ là "0000100000001010101", và tại lối vào bộ định tuyến 17 sẽ là "0000000010000000101", hướng đến cổng phía Đông đã bị chặn. Do đó, hoạt động định tuyến cần thay đổi thông tin định tuyến phù hợp để bảo đảm giao tác truyền thông hướng đến đích tin cậy, có thể cập nhật thông tin định tuyến tại bộ định tuyến 17 để đường định tuyến chuyển qua các bộ định tuyến {12,11,10,5}, đường nét đứt trong Hình 3.1. Có nghĩa là thông tin của trường PTT tại bộ định tuyến 17 cần cập nhật mới là "0000000010000010100".

- Hoạt động truyền thông của mạng trên chip sử dụng kỹ thuật chuyển mạch gói, gói tin được chia nhỏ thành các flit 34 *bit* với cấu trúc được mô tả trong Hình 3.3. Trong đó, thông tin định tuyến được thiết lập sẵn tại trường PTT của flit tiêu đề để xác định đường đi của thông tin trong mạng [36].



Hình 3.3: Cấu trúc gói tin.

Theo nguyên lý điều chuyển dữ liệu *wormhole* và định tuyến tại nguồn, tại mỗi bộ định tuyến thông tin định tuyến có sẵn trong trường PTT được sử dụng để xác định hướng đi tiếp theo của gói tin. Hoạt động chuyển tiếp thông tin của bộ định tuyến bắt đầu khi cổng lối vào nhận được flit tiêu đề và trích xuất 2 *bit* trọng số thấp nhất của trường PTT để xác định cổng lối ra. Dựa trên trên thông tin định tuyến, cổng lối vào sẽ tạo tín hiệu yêu cầu chiếm tài nguyên của cổng lối ra; tín hiệu chiếm tài nguyên này sẽ bị xóa khi nhận được một flit cuối cùng (*tail flit*) của gói tin. Khối cổng lối ra dựa vào các tín hiệu yêu cầu chiếm tài nguyên để phân xử cấp quyền chuyển tiếp thông tin cho các khối cổng lối vào và giải phóng tài nguyên khi tín hiệu này bị xóa. Trường định tuyến PTT cũng được dịch sang phải 2 *bit* để sẵn sàng cho việc định tuyến tại bộ định tuyến tiếp theo (có nghĩa là giá trị 2 *bit* trọng số thấp cũ bị xóa và thay thế bằng giá trị mới).

Trên cơ sở phân chia các đoạn của một đường định tuyến (Mục 3.1.1), để đảm bảo hoạt động cập nhật đường định tuyến một cách hiệu quả chúng ta cần

phải xác định vị trí tương đối của bộ định tuyến bị cấm. Từ việc xác định vị trí của bộ định tuyến bị cấm sẽ cho phép thực hiện các giải thuật thay đổi đường định tuyến phù hợp với từng trường hợp cụ thể. Nguyên lý xác định vị trí bộ định tuyến bị cấm được hình thành dựa trên sự kết hợp giữa đặc điểm xác định đường định tuyến của thuật toán XY và nguyên tắc thiết lập thông tin định tuyến bằng trường định tuyến PTT. Sau đây là một số nguyên tắc dùng cho việc xác định vị trí bộ định tuyến bị cấm:

1. Mỗi khối cổng vào/ra của bộ định tuyến không tự nhận biết được định danh của chính nó và không xử lý thay đổi thông tin của các flit đến và đi.

2. Thông tin định tuyến có tính chất tương đối và phụ thuộc nguyên tắc định danh cho các khối cổng bên trong mỗi bộ định tuyến của mạng trên chip.

3. Trong hoạt động truyền thông sử dụng giải thuật định tuyến tĩnh XY, mỗi đường định tuyến chỉ chuyển hướng tại một vị trí (góc định tuyến) và chuyển hướng đến cổng cục bộ tại bộ định tuyến đích đến.

4. Trên đoạn thẳng định tuyến, thông tin định tuyến tương ứng của mỗi bộ định tuyến là như nhau do mỗi bộ định tuyến nhận thông tin cùng một cổng lối vào và chuyển tiếp đến lối ra cùng phương. Do đó, tại mỗi bộ định tuyến giá trị thông tin định tuyến của bộ định tuyến hiện tại và ít nhất hai bộ định tuyến liền sau là như nhau.

5. Tại góc định tuyến, thông tin định tuyến hướng đến cổng lối ra của một phương khác với cổng lối vào và cũng khác với thông tin định tuyến của bộ định tuyến liền trước.

6. Trong trường hợp đặc biệt, khi bộ định tuyến tại đích đến của giao tác truyền thông nằm trong số các bộ định tuyến bị ảnh hưởng của hoạt động thay đổi đường truyền hoặc ngay sau bộ định tuyến mới được thay đổi. Lúc này, thông tin định tuyến cần được thay đổi giá trị bằng với định danh cổng lối vào mới để phù hợp với kiến trúc của bộ định tuyến của mạng trên chip [34].

Để thay đổi đường định tuyến, thông tin định tuyến cần được cập nhật tại bộ định tuyến ngay trước vị trí bộ định tuyến bị cấm và cần đảm bảo một số yêu cầu sau:

- Hoạt động cập nhật định tuyến chỉ thực hiện tại bộ định tuyến ngay trước bộ định tuyến bị cấm trên đường định tuyến.

- Đường định tuyến mới là sự thay đổi ít nhất nhưng vẫn bảo đảm thông tin hướng đến đích tin cậy.
- Kết hợp sử dụng giải thuật định tuyến góc mở Phía Tây Trước (WF) [23].
- Phù hợp với đặc điểm hoạt động của bộ định tuyến cho mạng trên chip [34].

3.2 Giải pháp cập nhật thông tin định tuyến

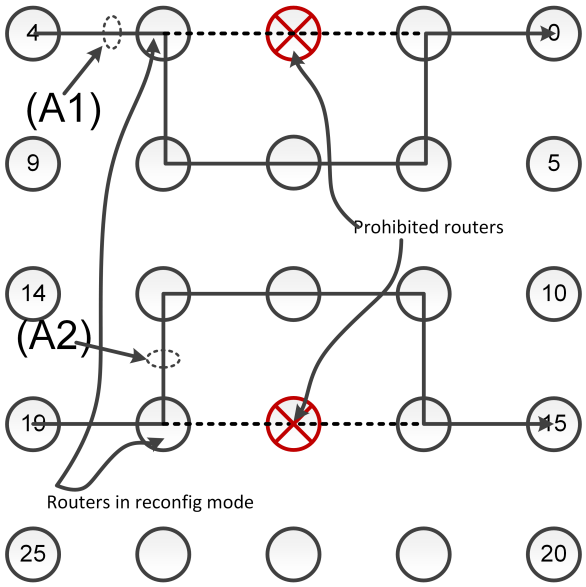
Giải pháp tái cấu hình được đề xuất cho phép mạng trên chip bảo đảm cung cấp các dịch vụ truyền thông tin cậy ngay cả khi có bộ định tuyến rời khỏi mạng, làm thay đổi cấu hình của mạng trên chip. Do đó, giải pháp cập nhật thông tin định tuyến được đề xuất dựa trên cơ sở phân tích thông tin định tuyến hiện tại nhằm bảo đảm hệ thống hoạt động ổn định. Dựa trên đặc điểm định tuyến của giải thuật định tuyến XY đã được phân tích, giải pháp cập nhật thông tin định tuyến được đề xuất phân các trường hợp cần thay đổi đường định tuyến thành ba nhóm sau đây:

1. Trường hợp A: Cập nhật định tuyến khi bộ định tuyến bị cấm nằm trên đoạn thẳng.
2. Trường hợp B: Cập nhật định tuyến khi bộ định tuyến bị cấm nằm tại góc định tuyến.
3. Trường hợp C: Cập nhật định tuyến khi bộ định tuyến bị cấm nằm ở lân cận góc định tuyến.

3.2.1 Cập nhật định tuyến khi bộ định tuyến bị cấm nằm trên đoạn thẳng định tuyến

Xét trên một đoạn thẳng định tuyến của giao tác truyền thông sử dụng giải thuật định tuyến tĩnh XY như đã đề cập trong Mục 3.1.1, giả thiết có một bộ định tuyến ở trên đoạn thẳng định tuyến đó đã rời khỏi mạng (*do bị lỗi, hỏng hóc hoặc không còn được sử dụng*) thì giao tác truyền thông sẽ bị chặn lại tại bộ định tuyến ngay trước bộ định tuyến bị cấm. Hoạt động tái cấu hình cần được thực hiện để thay đổi đường định tuyến phù hợp với cấu hình mạng hiện thời tại bộ định tuyến tái cấu hình. Hình 3.4, minh họa giải pháp cập nhật định tuyến với trường hợp cập nhật định tuyến khi bộ định tuyến bị cấm nằm trên đường thẳng. Trong đó, đường định tuyến trước khi thay đổi là *đường đứt nét*

và đường định tuyến sau khi đã thay đổi thông qua hoạt động tái cấu hình là *đường liền nét*.



Hình 3.4: Cập nhật định tuyến khi bộ định tuyến bị cấm nằm trên đoạn thẳng định tuyến (trường hợp A).

Để đảm bảo hoạt động cập nhật đường định tuyến với ít thay đổi nhất, giải pháp thay đổi đường định tuyến bằng cách truyền thông vòng qua bộ định tuyến bị cấm và kết hợp với nguyên lý chọn hướng của giải thuật định tuyến Phía Tây trước. Đường định tuyến đi vòng sẽ bắt đầu từ bộ định tuyến tái cấu hình (nằm ngay trước bộ định tuyến bị cấm) và hướng trở về đường định tuyến cũ ở vị trí bộ định tuyến thứ 3 ngay sau bộ định tuyến tái cấu hình (hay nói cách khác là quay trở lại đường định tuyến cũ tại vị trí sớm nhất có thể). Để thực hiện được sự thay đổi đường định tuyến của giao tác truyền thông như trong Hình 3.4, bộ định tuyến tái cấu hình cần thay đổi nội dung trường định tuyến PTT như mô tả trong Hình 3.5. Nội dung trường định tuyến PTT ở Hình 3.5 thể hiện hoạt động cập nhật định tuyến cần bổ sung góc rẽ tại bộ định tuyến tái cấu hình (2 *bit* trọng số thấp) và thông tin định tuyến cho phép trở lại đường định tuyến cũ (ở các bit thứ 7 và 8 từ phải sang).

PTT bị chặn	xx	xx	xx	xx	xx	xx	01	01	01	
PTT đã cập nhật	xx	xx	xx	xx	01	00	01	01	10	Giao tác (A1)
PTT đã cập nhật	xx	xx	xx	xx	01	10	01	01	00	Giao tác (A2)

Hình 3.5: Thông tin trường định tuyến tương ứng với Hình 3.4.

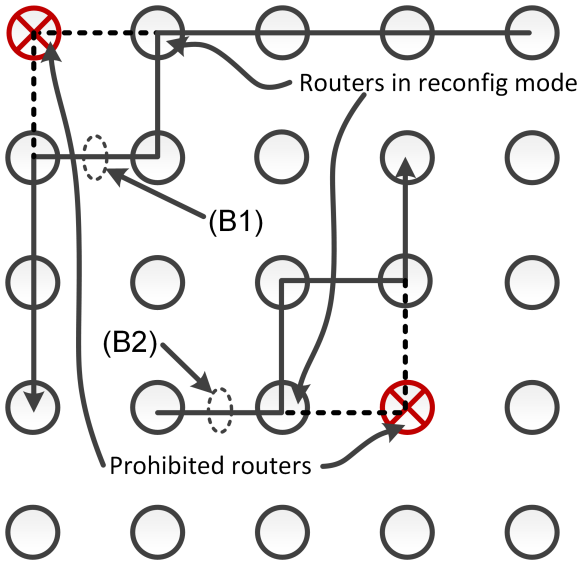
Các bit có ký hiệu "xx" có nghĩa là giá trị bất kỳ hay không quan tâm

trong xem xét này [1].

Trong trường hợp đi vòng qua bộ định tuyến bị chặn, hoạt động cập nhật định tuyến sẽ làm tăng thêm hai nút mạng so với đường định tuyến cũ. Để bảo đảm hoạt động cập nhật định tuyến không làm xuất hiện hiện tượng khóa chết trong hoạt động truyền thông mạng, giải thuật cập nhật định tuyến sẽ ưu tiên chọn hướng theo giải thuật định tuyến Phía Tây trước (xem Mục 1.4.3).

3.2.2 Cập nhật định tuyến khi bộ định tuyến bị cấm nằm tại góc định tuyến

Khi bộ định tuyến bị cấm được xác định là ở ngay vị trị góc định tuyến của một giao tác truyền thông đang xảy ra thì hoạt động cập nhật định tuyến cần thay đổi vị trí góc định tuyến phù hợp với cấu hình hiện thời của mạng. Hình 3.6 mô tả một số ví dụ cho trường hợp cập nhật định tuyến khi bộ định tuyến bị cấm nằm tại góc định tuyến (B). Trong đó, đường định tuyến cũ được minh họa bởi đường đứt nét và đường định tuyến sau khi đã thay đổi là đường liền nét. Do đó, bộ định tuyến tái cấu hình (là bộ định tuyến ngay trước bộ định tuyến bị cấm) cần xác định được vị trí tương đối của bộ định tuyến bị cấm và hướng rẽ của góc định tuyến để đưa ra quyết định cập nhật định tuyến phù hợp.

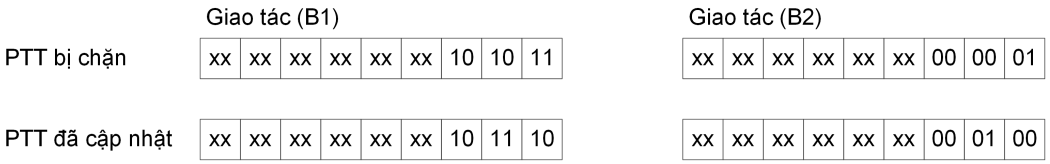


Hình 3.6: Cập nhật định tuyến khi bộ định tuyến bị cấm nằm tại góc định tuyến (trường hợp B).

Từ kết quả phân tích đặc điểm của giải thuật định tuyến tĩnh XY và yêu cầu hạn chế thay đổi đường định tuyến như đã trình bày trong Mục 3.1, hoạt

[1]Ký hiệu 'x' tương ứng với "don't care" trong quy ước của IEEE 1161 Standard

động tái cấu hình để cập nhật định tuyến sẽ đưa thay đổi góc định tuyến hướng vào phía trong của đường định tuyến trước đó. Hình 3.7 mô tả nội dung cần thay đổi cho trường định tuyến PTT tương ứng với trường hợp cập nhật tại góc định tuyến đã mô tả trong Hình 3.6. Trong trường hợp cập nhật tại góc định tuyến, hoạt động cập nhật định tuyến chỉ hướng đến sự thay đổi vị trí góc định tuyến nên không làm tăng số lượng nút mạng trung gian.



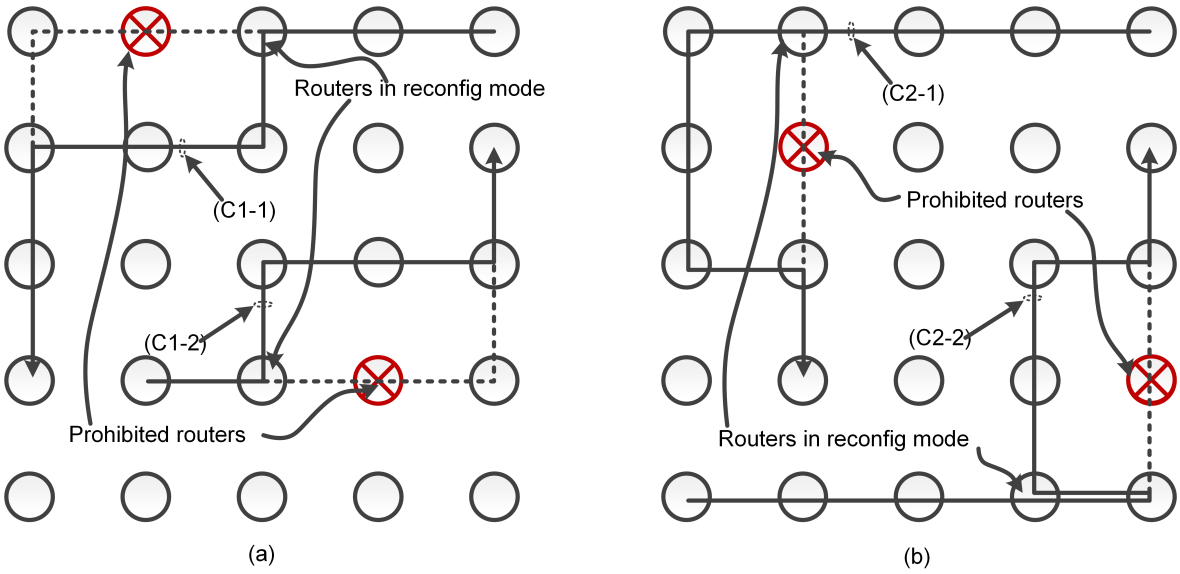
Hình 3.7: Thông tin trường định tuyến tương ứng với Hình 3.6.

Hoạt động thay đổi góc định tuyến cần chú ý hơn đối với các giao tác truyền thông có đích đến ngay sau góc định tuyến, nhưng trong hoạt động tái cấu hình tại góc định tuyến thì hướng vào của bộ định tuyến đích đến đã bị thay đổi. Đồng thời, Hình 3.2 mô tả nguyên tắc liên kết giữa cổng cục bộ được kết nối với các cổng vào/ra khác ở các hướng được xác định bằng chính định danh của cổng đó. Do vậy, hoạt động cập nhật thông tin định tuyến cần xử lý thay đổi thông tin định tuyến đích đến phù hợp với cổng lối vào mới. Thông tin định tuyến đích đến sẽ được thay đổi bằng cách lấy bù $2bit$ trọng số cao của thông tin định tuyến ở vị trí góc định tuyến mới.

3.2.3 Cập nhật định tuyến khi bộ định tuyến bị cấm nằm ở lân cận góc định tuyến

Khi vị trí của bộ định tuyến bị cấm ở lân cận góc định tuyến của giao tác truyền thông, hoạt động cập nhật định tuyến cần xử lý kết hợp nguyên tắc của cả hai trường hợp trên (xem Mục 3.2.1 và Mục 3.2.2). Trong Hình 3.8 mô tả các trường hợp cập nhật định tuyến khi bộ định tuyến bị cấm ở lân cận góc định tuyến. Để minh họa hoạt động cập nhật định tuyến tường minh hơn, giải thuật cập nhật định tuyến chia làm hai trường hợp con: bộ định tuyến bị cấm nằm ngay trước góc định tuyến (C1) và bộ định tuyến bị cấm nằm ngay sau góc định tuyến (C2) để có những thay đổi đường định tuyến phù hợp.

Dựa trên đặc điểm kiến trúc bộ định tuyến của mạng trên chip sử dụng giải thuật định tuyến tĩnh XY, giải pháp cập nhật cho bộ định tuyến bị cấm ở liền trước góc định tuyến thay đổi luôn góc định tuyến ngay tại vị trí của bộ định tuyến tái cấu hình (bộ định tuyến nằm ngay trước bộ định tuyến bị cấm).



Hình 3.8: Cập nhật định tuyến khi bộ định tuyến bị cấm ở lân cận góc định tuyến (trường hợp C).

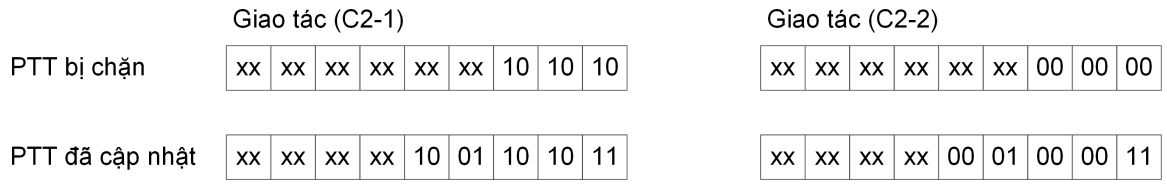
Do đó, trường hợp này hoạt động cập nhật định tuyến không làm thay đổi số lượng nút mạng trung gian của giao tác truyền thông. Trong trường hợp đặc biệt, bộ định tuyến đích đến ở ngay sau góc định tuyến thì giá trị định tuyến tại vị trí này cần được thay đổi phù hợp với hướng đến mới của nó để hướng gói tin đến cổng cục bộ. Để minh họa cho hoạt động cập nhật thông tin định tuyến, Hình 3.9 biểu diễn hoạt động thay đổi nội dụng thông tin định tuyến của trường PTT trong trường hợp tương ứng với Hình 3.8a.

	Giao tác (C1-1)									Giao tác (C1-2)								
PTT bị chặn	xx	xx	xx	xx	xx	10	10	11	11	xx	xx	xx	xx	xx	00	00	01	01
PTT đã cập nhật	xx	xx	xx	xx	10	01	11	11	10	xx	xx	xx	xx	xx	00	01	01	00

Hình 3.9: Thông tin trường định tuyến tương ứng với Hình 3.8a.

Đối với trường hợp bộ định tuyến bị cấm nằm ngay sau góc rẽ định tuyến (Hình 3.8b) thì chỉ bộ định tuyến tại góc rẽ đường định tuyến mới xác định hướng và thực hiện hoạt động tái cấu hình. Bộ định tuyến tái cấu hình (bộ định tuyến tại góc định tuyến) sẽ ưu tiên chuyển tiếp thông tin vòng qua bộ định tuyến bị cấm theo hướng mới. Trong trường hợp đặc biệt, khi bộ định tuyến bị cấm nằm ở biên của mạng, bộ định tuyến tái cấu hình phải chuyển ngược gói tin theo hướng của lối cổng lối vào. Điều này chỉ có thể xảy ra khi bộ định tuyến hoạt động ở chế độ tái cấu hình. Để minh họa rõ hơn hoạt động thay đổi đường định tuyến của giải pháp cập nhật thông tin định tuyến, nội dung thông tin định tuyến của trường PTT thay đổi tương ứng với hai giao tắc ở trường

hợp của Hình 3.8b được trình bày trong Hình 3.10.



Hình 3.10: Thông tin trường định tuyến tương ứng với Hình 3.8b.

3.3 Kiến trúc bộ định tuyến tái cấu hình đề xuất

Để thực hiện giải pháp tái cấu hình đề xuất tại Mục 3.2, mục này trình bày kiến trúc bộ định tuyến tái cấu hình cho phép thay đổi thông tin định tuyến. Theo đó, bộ định tuyến đề xuất này có thể hoạt động ở hai chế độ: chế độ bình thường (*normal mode*) và chế độ tái cấu hình (*reconfig mode*).

Bộ định tuyến hoạt động trong **chế độ bình thường (*normal mode*)** chuyển tiếp gói tin của giao tác truyền thông theo hướng định tuyến đã xác định từ trước. Ở chế độ bình thường, gói tin nhận được từ cổng lối vào sẽ được chuyển hướng đến lối ra một cách lần lượt, từ flit tiêu đề đến khi flit cuối cùng được chuyển tới cùng một cổng lối ra được chọn. Chế độ bình thường được thiết lập mặc định trong hoạt động chuyển tiếp thông tin của mọi bộ định tuyến trong mạng. Một bộ định tuyến luôn hoạt động ở chế độ bình thường khi bốn bộ định tuyến lân cận của nó không bị cấm; và ngay cả những bộ định tuyến có hướng ra bị chặn vẫn hoạt động ở chế độ bình thường đối với các giao tác truyền thông chuyển tiếp qua nó mà không được hướng đến lối ra đã bị chặn.

Chế độ tái cấu hình (*reconfig mode*) của bộ định tuyến được thiết lập khi một gói tin nhận được ở đầu vào của bộ định tuyến có đường định tuyến hướng đến bộ định tuyến bị cấm liền ngay sau nó. Lúc này bộ định tuyến cần hoạt động ở chế độ tái cấu hình để xử lý và cập nhật thông tin định tuyến để thích ứng với sự thay đổi cấu hình và đáp ứng tối ưu các yêu cầu truyền thông của hệ thống. Trong chế độ tái cấu hình, hoạt động chuyển tiếp thông tin vẫn phải bảo đảm các quy tắc chiếm tài nguyên và phân xử nhằm ổn định hoạt động truyền thông của toàn mạng trên chip.

3.3.1 Giải pháp kiến trúc cho bộ định tuyến

Kiến trúc bộ định tuyến đề xuất sẽ được trình bày chi tiết ở các mục con sau đây. Kiến trúc bộ định tuyến đề xuất cần đáp ứng hoạt động trong mạng

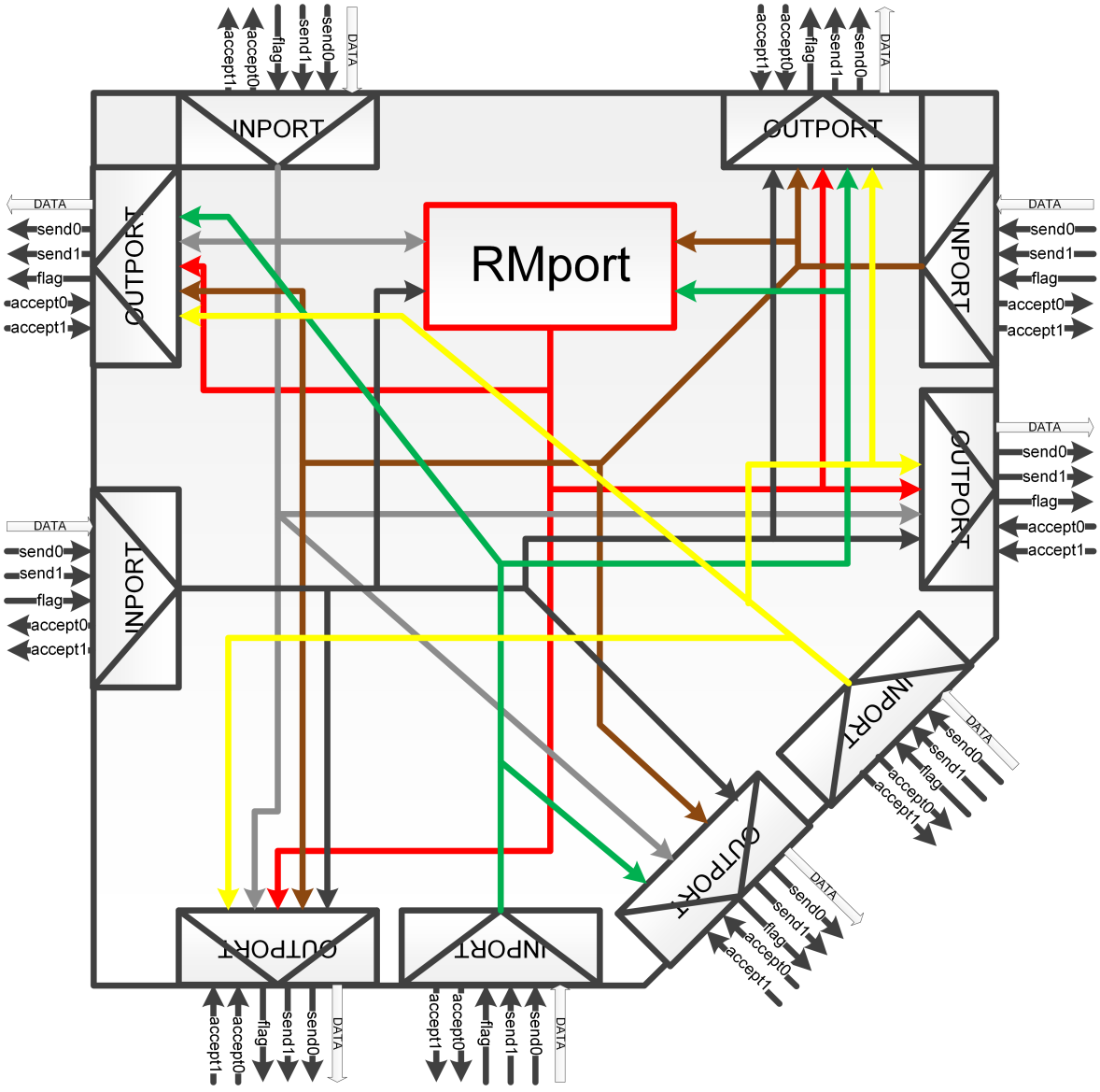
trên chip với một số đặc điểm sau:

- Mạng có cấu trúc dạng lưới hai chiều, sử dụng cơ chế điều chuyển dữ liệu WH và cơ chế điều khiển luồng *credit-based*.
- Thực thi hoạt động tái cấu hình theo nguyên lý xử lý phân tán để không phát sinh yêu cầu quảng bá thông tin cấu hình.
- Hoạt động ở hai chế độ: chế độ bình thường và chế độ tái cấu hình.
- Hoạt động tái cấu hình được thực hiện độc lập, hạn chế ảnh hưởng đến các hoạt động truyền thông ổn định khác.

Hình 3.11 mô tả kiến trúc đề xuất cho bộ định tuyến mạng trên chip tái cấu hình; với bốn khối cổng vào/ra theo bốn hướng (**Bắc, Đông, Nam và Tây**), một cổng cục bộ (**Local port**) dùng để kết nối lõi IP với mạng và cổng ảo thực thi tái cấu hình (**RMport: *Routing Modification port***) để thực hiện hoạt động cập nhật thông tin định tuyến. Với định hướng đưa ra giải pháp tái cấu hình dựa trên nguyên lý xử lý phân tán, trạng thái hoạt động của bộ định tuyến không phát quảng bá mà được thiết lập trực tiếp qua các cờ trạng thái. Do đó, một tín hiệu cờ trạng thái (***flag***) được sử dụng để thông báo trạng thái cấm của bộ định tuyến bị cấm (hướng từ cổng lối vào của bộ định tuyến bị cấm đến bốn bộ định tuyến lân cận). Từ trạng thái cờ nhận được các cổng lối ra sẽ thiết lập trạng thái hoạt động cho chính nó và lập cờ bị chặn tương ứng đến tất cả các cổng lối vào trong cùng bộ định tuyến và khối cổng ảo RMport. Hoạt động chuyển tiếp trạng thái này cho phép các khối cổng lối vào nhận biết trạng thái của hướng chuyển tiếp gói tin và có thể tự chuyển vào chế độ tái cấu hình khi đường định tuyến bị chặn; đồng thời các cờ trạng thái này cũng cho phép ***RMport*** xác định hướng chuyển tiếp thông tin phù hợp với cấu hình mạng hiện thời.

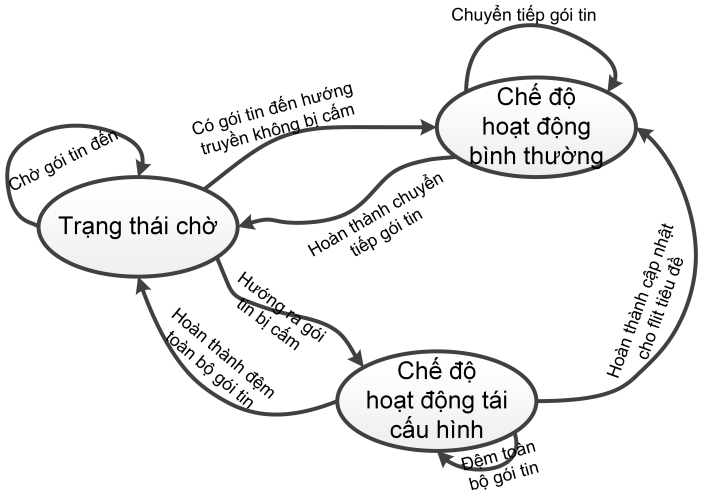
Để hạn chế một cách tối đa các yêu cầu tính toán và xử lý trong quá trình cập nhật định tuyến, kiến trúc đề xuất không thực hiện tái cấu hình tại các bộ định tuyến được nối trực tiếp với nguồn tin. Do vậy, kiến trúc đề xuất không liên kết khối cổng ảo cập nhật thông tin định tuyến với khối cổng vào/ra cục bộ. Thay vào đó, khối cổng lối ra sẽ chuyển tiếp trạng thái của bốn hướng lân cận cho cổng cục bộ để khối phát tin làm cơ sở tự xác định đường định tuyến phù hợp với cấu hình hiện thời của mạng.

Nguyên lý hoạt động của bộ định tuyến có thể tái cấu hình RNoC được minh họa ở dạng máy trạng thái hữu hạn như trong Hình 3.12; với ba trạng thái



Hình 3.11: Kiến trúc bộ định tuyến có thể tái cấu hình đề xuất.

cơ bản: trạng thái chờ (*idle*), chế độ hoạt động bình thường (*normal mode*) và chế độ hoạt động tái cấu hình (*reconfig mode*).



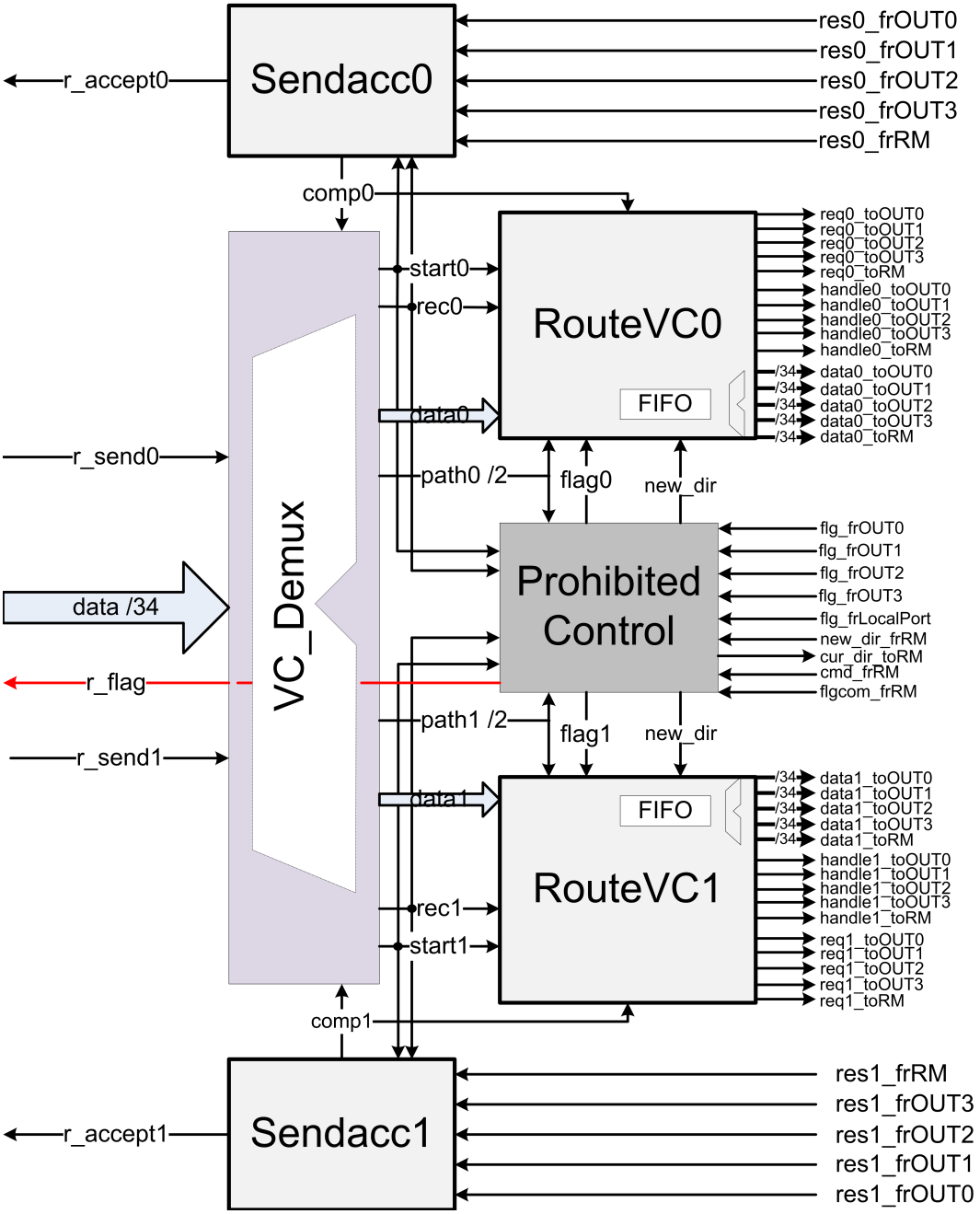
Hình 3.12: Cơ chế chuyển trạng thái của bộ định tuyến RNoC.

Hoạt động của bộ định tuyến RNoC dựa trên nguyên lý trao đổi thông tin đồng bộ theo xung nhịp, sau khi được khởi động hoặc tín hiệu *reset* từ mức tích cực âm (*logic 0*) chuyển sang mức logic 1, bộ định tuyến ở trạng thái sẵn sàng nhận tin. Khi có gói tin được gửi đến, khối cổng lối vào nhận được flit tiêu đề sẽ trích thông tin định tuyến từ trường PTT để xác định hướng ra được chọn. Nếu hướng ra không bị cấm (không kết nối với bộ định tuyến bị cấm), toàn bộ gói tin sẽ được chuyển đến cổng lối ra tương ứng và bộ định tuyến sẽ hoạt động ở chế độ hoạt động bình thường. Trong trường hợp hướng ra bị cấm, khối cổng lối vào sẽ chuyển sang hoạt động ở chế độ tái cấu hình (*reconfig mode*) và flit tiêu đề được gửi đến khối cập nhật định tuyến để xác định hướng ra mới và cập nhật thông tin định tuyến; lúc này bộ định tuyến hoạt động ở chế độ tái cấu hình. Dựa vào giải pháp cập nhật định tuyến đã đề cập ở Mục 3.2, khối cập nhật định tuyến kiểm tra các hướng ra khả dĩ và thực hiện lựa chọn lối ra cùng lúc với việc cập nhật lại thông tin ở trường định tuyến của flit tiêu đề. Với cách thực hiện này, chỉ có flit tiêu đề được chuyển tới khối cổng ảo còn các flit tiếp theo sẽ được định tuyến trực tiếp đến lối ra mới. Trong trường hợp đặc biệt, gói tin cần được định tuyến quay trở về hướng trùng với hướng đến; nhưng đo đặc điểm kiến trúc của bộ định tuyến mạng trên chip thì khối cổng lối vào không thể chuyển tin đến cổng lối ra cùng hướng (Mục 1.2). Do vậy, toàn bộ gói tin sẽ được chuyển qua khối cổng ảo cập nhật định tuyến trước khi đưa ra lối ra cùng hướng với cổng lối vào.

3.3.2 Kiến trúc chi tiết khối cổng lối vào và lối ra

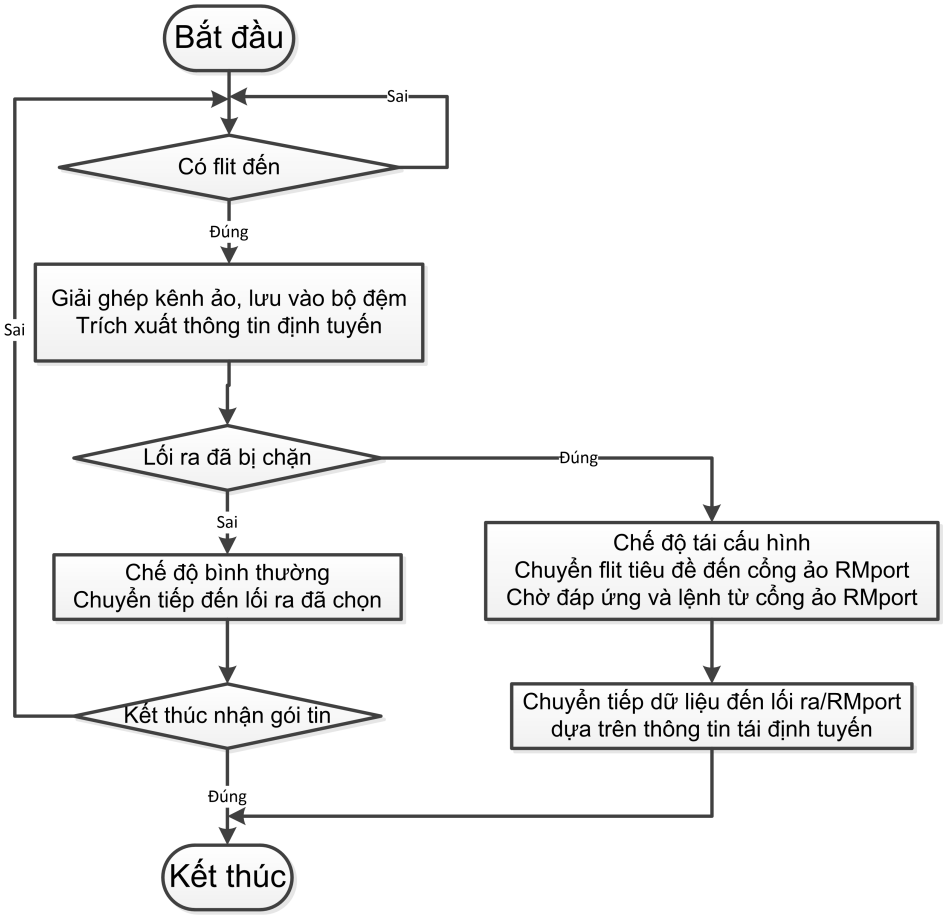
Thực hiện chức năng nhận thông tin đầu vào cho bộ định tuyến là khối cổng lối vào (***INPORT***), có các thành phần được mô tả chi tiết như trong Hình 3.13. Trong đó khối phân tách kênh ảo đầu vào (***VC-Demux***) có chức năng tách kênh ảo từ liên kết vật lý và hướng đến bộ đệm đầu vào của kênh ảo tương ứng, đồng thời khối này còn có chức năng trích xuất thông tin định tuyến từ *flit* tiêu đề để lựa chọn lối ra theo đường định tuyến đã được xác lập tại nguồn tin. Khối điều khiển cấm (***Prohibited Control***) có chức năng phát hiện hướng truyền bị cấm để chuyển hoạt động của khối cổng lối vào sang chế độ tái cấu hình, đồng thời thực hiện chức năng điều khiển bắt tay với khối cổng ảo cập nhật định tuyến để thực hiện chuyển tiếp các flit còn lại của gói tin tương ứng với các trường hợp cập nhật định tuyến cụ thể. Hai khối định tuyến kênh ảo ***RouteVC0*** & ***RouteVC1*** thực hiện chức năng lưu trữ và chuyển tiếp thông tin theo sự điều khiển của khối con điều khiển; và hai khối ***Sendacc0&Sendacc1*** có chức năng trao đổi tín hiệu bắt tay với khối cổng lối ra và với bộ định tuyến liền trước tương ứng với kênh ảo 0 hoặc 1. Khi bộ định tuyến được chỉ định bị cấm, trạng thái cấm được thiết lập đến các khối cổng lối vào thông qua khối cổng lối ra cục bộ (*Local Outport*); lúc này khối con điều khiển sẽ thiết lập cờ trạng thái ở mức tích cực cao (logic 1).

Hoạt động của khối cổng lối vào khi nhận một gói tin hoàn thiện được thể hiện như lưu đồ ở Hình 3.14. Khối cổng lối vào luôn ở trạng thái chờ sau khi hệ thống khởi động để sẵn sàng đáp ứng những yêu cầu chuyển tiếp thông tin. Khi có tín hiệu yêu cầu nhận tin tác động đến lối vào, khối cổng lối vào sẽ tách kênh ảo và trích xuất thông tin định tuyến để lưu giữ và hướng dữ liệu đến khối cổng lối ra phù hợp. Trong trường hợp *flit* tiêu đề, thông tin định tuyến được trích xuất để xác định hướng ra theo nguyên lý định tuyến và cũng để kiểm tra trạng thái của lối ra tại khối điều khiển ***Prohibited Control***. Dữ liệu 34 *bit* từ giao tiếp lối vào sẽ được tách kênh ảo để lưu và bộ đệm tương ứng với kênh ảo đầu vào tại một trong hai khối ***RouteVC0/RouteVC1***. Khối điều khiển ***Prohibited Control*** dựa vào trạng thái của các cổng lối ra và thông tin định hướng để nhận biết hướng đi của gói tin có bị cấm hay có thể chuyển tiếp. Trong trường hợp không bị cấm, khối điều khiển không hoạt động, khối cổng lối vào hoạt động ở chế độ bình thường và gói tin được chuyển tới khối cổng lối ra được xác định bởi thông tin định tuyến. Khi một gói tin lối ra hướng đến bộ định tuyến bị chặn, khối điều khiển sẽ lập cờ trạng thái và chuyển khối cổng lối vào sang chế độ tái cấu hình, thông tin định tuyến được chuyển đến khối cổng ảo cập nhật



Hình 3.13: Kiến trúc chi tiết của khối cổng lối vào.

định tuyến. Lúc đó các khối con khác của khối cổng lối vào được điều khiển để liên kết với với khối cổng ảo cập nhật định tuyến. Hoạt động chuyển tiếp thông tin được thực hiện ở hai khối con ***RouteVC0***&***RouteVC1*** tương ứng với hai kênh ảo được sử dụng của mạng trên chip. Chức năng thiết lập và trao đổi các tín hiệu bắt tay được thực hiện tại hai khối con ***Sendacc0***&***Sendacc1*** phù hợp với kênh ảo nhận tin. Các khối con này hoạt động dưới sự điều khiển của khối con ***Prohibited Control*** trong chế độ tái cấu hình, và liên kết trực tiếp với các khối cổng lối ra khác trong chế độ hoạt động bình thường của bộ định tuyến.



Hình 3.14: Lưu đồ hoạt động của khối cổng lối vào.

Khối cổng lối ra (***OUTPORT***) có chức năng chuyển tiếp thông tin từ các cổng lối vào hoặc khối cổng ảo cập nhật định tuyến theo sự điều khiển và phân xử lối vào với các khối chức năng và tín hiệu được mô tả chi tiết như trong Hình 3.15. Khối cổng lối ra có chức năng phân xử các yêu cầu chuyển tiếp thông tin từ các khối cổng lối vào và khối cổng ảo ***RMport*** bởi hai khối con ***ArbiterVC0***&***ArbiterVC1*** và sau đó thông tin sẽ được hướng đến bộ định tuyến liền sau thông qua các khối ghép kênh (***MuxVC0/MuxVC1*** và ***VC_Mux***). Khối điều khiển ***Prohibited Control*** có chức năng nhận biết

trạng thái của bộ định tuyến lân cận và truyền về khối cổng lối vào và khối cổng ảo ***RMport*** định tuyến trong trường hợp hướng truyền đã bị cấm.

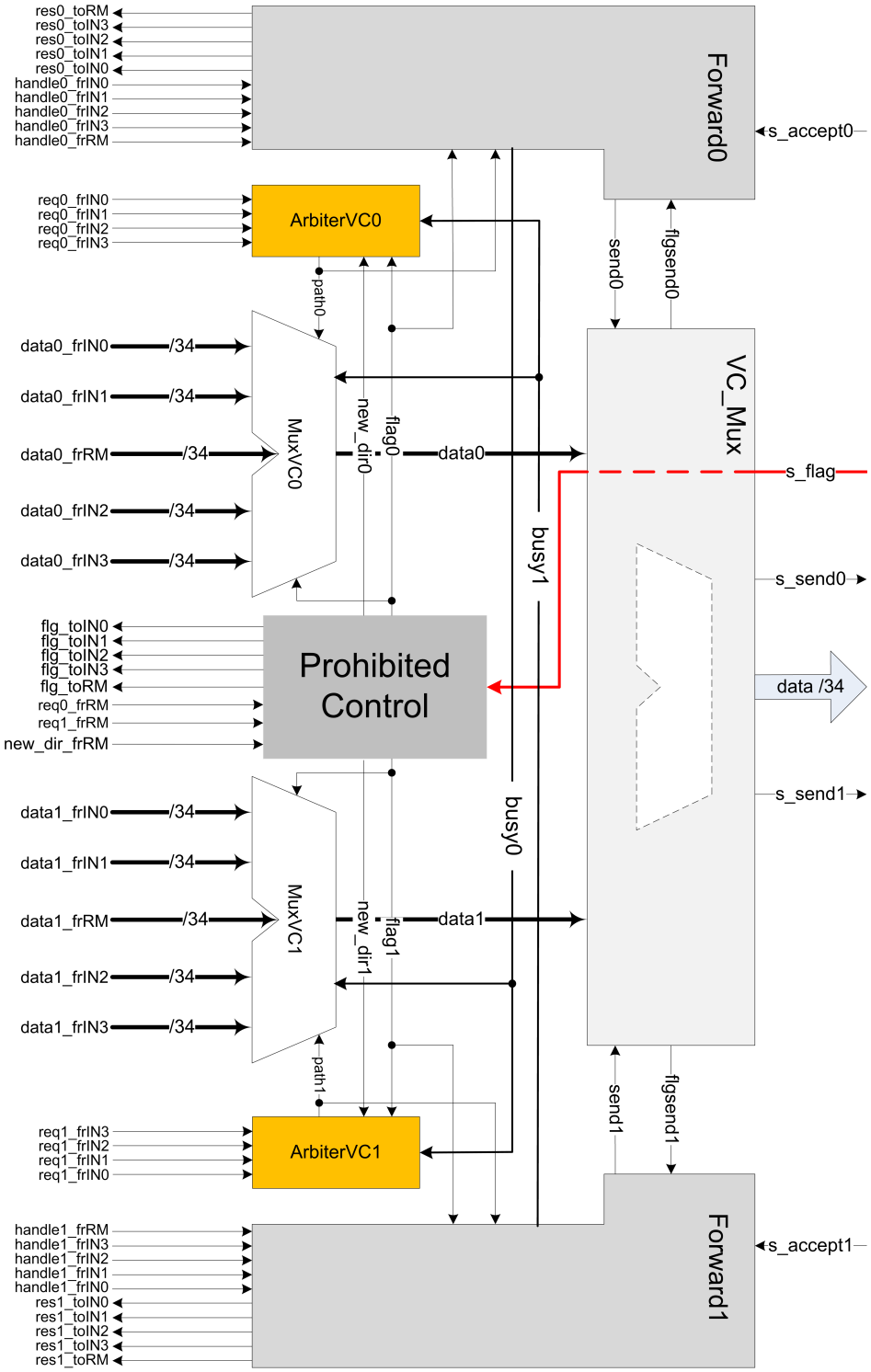
Nguyên lý hoạt động của khối cổng lối ra cho quá trình đáp ứng chuyển tiếp một gói tin được thể hiện qua lưu đồ như trong Hình 3.16. Ở chế độ chờ, khối cổng lối ra ở trạng thái sẵn sàng nhưng không tích cực các trạng thái phản hồi cho các khối cổng lối vào hoặc khối cổng ảo RMport. Trạng thái sẵn sàng đáp ứng chỉ chuyển sang trạng thái tích cực sau khi khối phân xử kênh ảo ***ArbiterVC0*** hoặc ***ArbiterVC1*** nhận được yêu cầu chuyển tiếp và quyết định đáp ứng yêu cầu cụ thể. Khi hoạt động ở chế độ bình thường (*normal mode*), khối cổng lối ra đáp ứng cho một yêu cầu từ các khối cổng lối vào và chuyển tiếp thông tin đến bộ định tuyến liền sau. Khi có yêu cầu từ khối cổng ảo ***RMport***, khối điều khiển ***Prohibited Control*** sẽ chờ đáp ứng hoàn tất giao dịch đang diễn ra để chuyển khối cổng lối ra vào chế độ tái cấu hình (*reconfig mode*). Ở chế độ tái cấu hình, khối cổng lối ra sẽ chuyển tiếp *flit* tiêu đề từ khối cổng ảo, sau đó sẽ dựa vào lệnh từ khối cổng ảo để quyết định duy trì chế độ cấu hình hoặc chuyển sang chế độ bình thường nhưng đáp ứng khối cổng lối ra do khối cổng ảo ***RMport*** chỉ định.

3.3.3 Kiến trúc thực hiện hoạt động cập nhật định tuyến

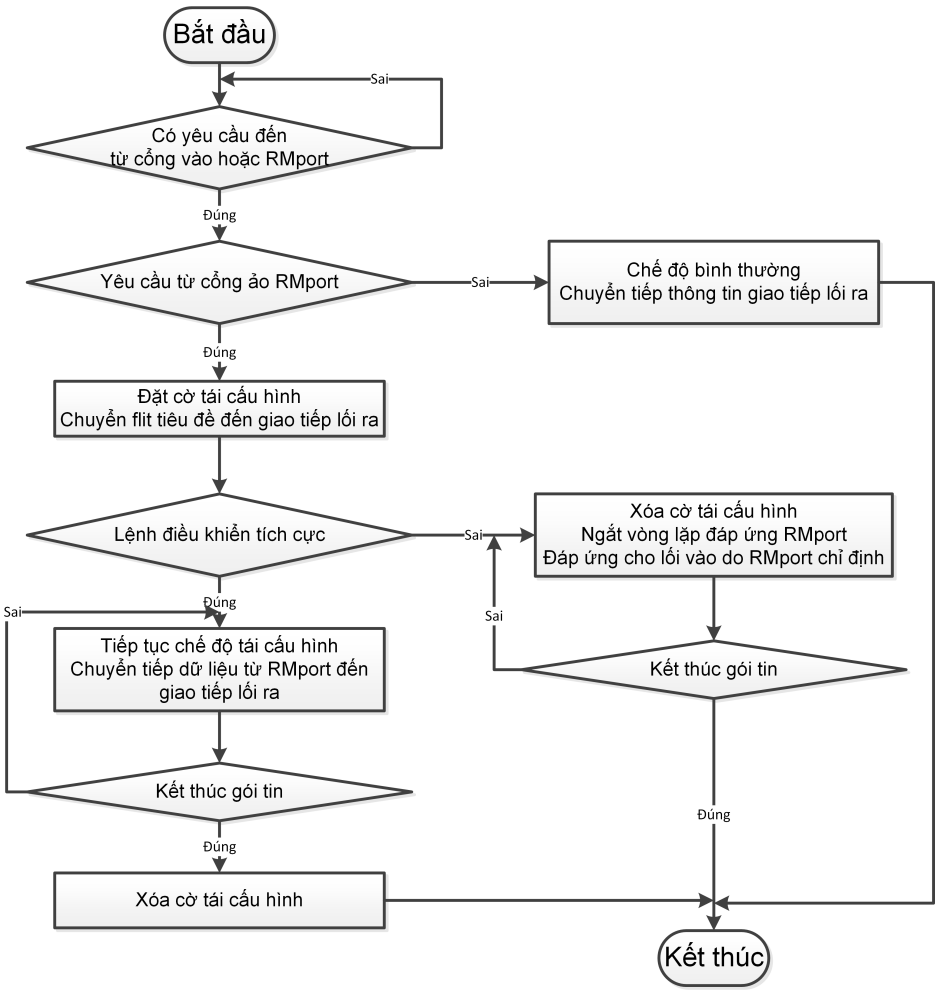
Một trong những hoạt động quan trọng của bộ định tuyến ở chế độ tái cấu hình là cập nhật đường định tuyến mới. Để thực hiện hoạt động cập nhật định tuyến tại mỗi bộ định tuyến được bổ sung khối cổng ảo thực thi tái cấu hình RMport. Khối ***RMport*** vừa có chức năng xử lý cập nhật định tuyến vừa có chức năng trao đổi dữ liệu với các khối cổng vào/ra nên được xem là một cổng ảo của bộ định tuyến. Hình 3.17 mô tả kiến trúc của cổng ảo ***RMport*** gồm các khối thành phần *Controller,Update, Receiver0, Receiver1, Send0* và *Send1*.

Để nhận được thông tin cần cập nhật cấu hình, khối cổng ảo ***RMport*** sử dụng hai khối ***Receiver0*** và ***Receiver1*** liên kết với cổng lối vào tương ứng cho hai kênh ảo 0 và 1 và sau đó thông tin được đưa qua bộ hợp kênh ảo để lưu vào bộ đệm nhận. Hoạt động nhận thông tin từ khối cổng lối vào tương đồng với hoạt động nhận tin của khối cổng lối ra của bộ định tuyến. Các *flit* sau khi được xử lý sẽ được chuyển hướng đến cổng lối ra thông qua hai khối ***Send0*** và ***Send1*** ứng với hai kênh ảo được hỗ trợ; yêu cầu chuyển tiếp thông tin này tương tự hoạt động chuyển tiếp thông tin từ khối cổng lối vào hướng đến khối cổng lối ra.

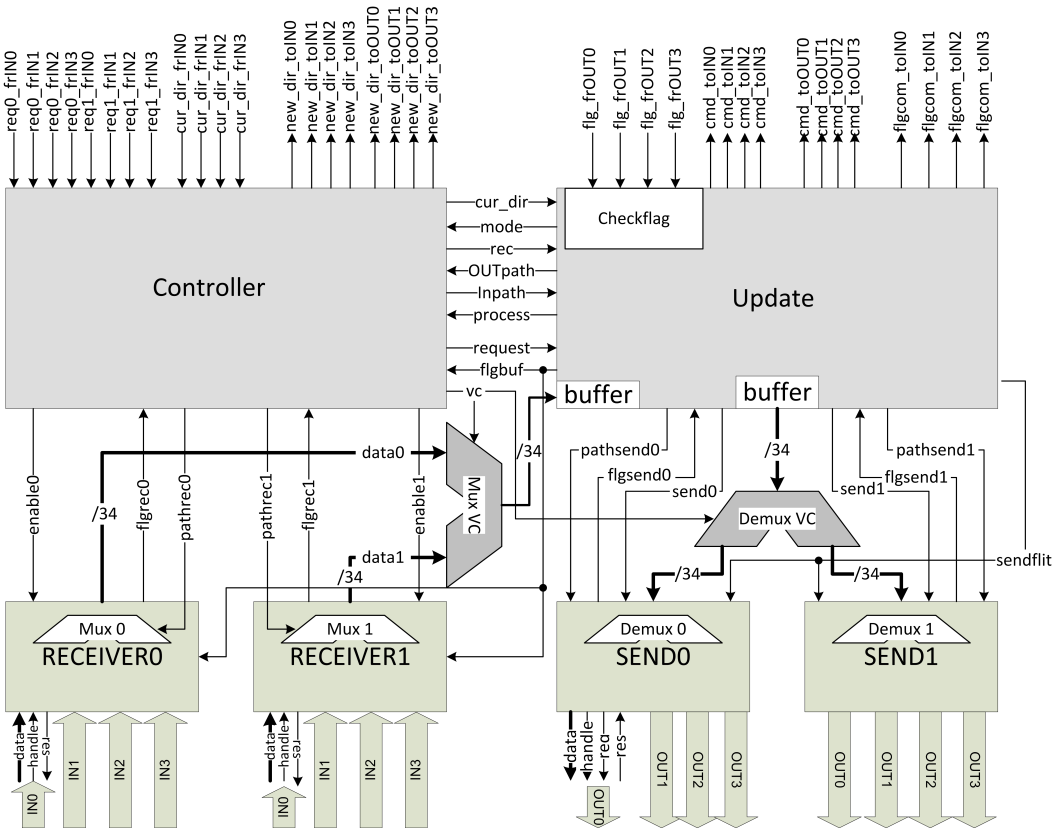
Khối ***Controller*** trong cổng ảo ***RMport*** thực hiện hai chức năng chính



Hình 3.15: Kiến trúc chi tiết của khối cổng lối ra.



Hình 3.16: Lưu đồ hoạt động của khối cổng lối ra.



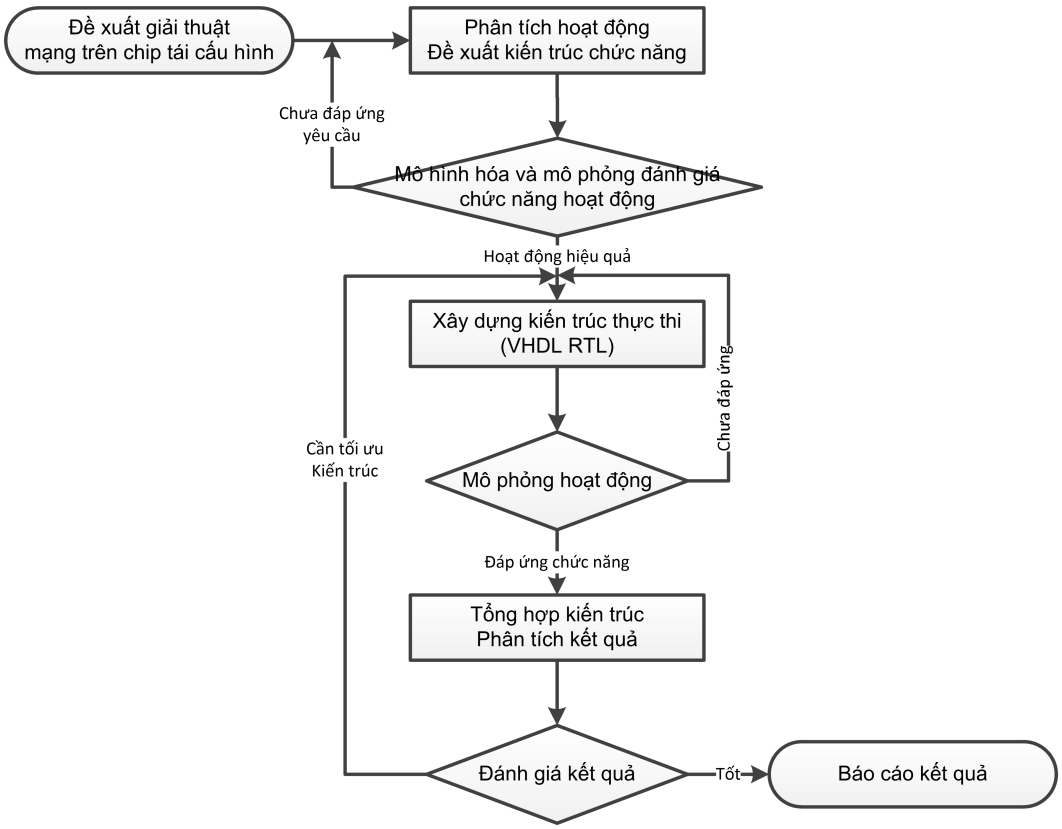
Hình 3.17: Kiến trúc chi tiết của khối cổng ảo thực thi tái cấu hình.

là điều khiển hoạt động của tất cả các khối con bên trong khối cổng ảo RMport, đồng thời trao đổi trạng thái và lệnh để điều khiển các khối cổng lối vào và cổng lối ra của bộ định tuyến khi hoạt động ở chế độ tái cấu hình. Bình thường, cổng ảo ***RMport*** ở chế độ chờ không hoạt động nếu không có yêu cầu tái cấu hình từ khối cổng lối vào. Khi có yêu cầu tái cấu hình gửi đến khối ***Controller***, khối con này sẽ thực hiện điều khiển khối nhận tin và khối cập nhật, sau đó chuyển tiếp thông tin đến khối cổng lối ra mới trong hoạt động tái cấu hình.

Khối cập nhật cấu hình ***Update*** thực hiện giải thuật thay đổi nội dung trường PTT của flit tiêu đề để cập nhật đường định tuyến mới của giao tác truyền thông. Tại khối ***Update***, nội dung thông tin định tuyến cũ sẽ được phân tích để xác định vị trí tương đối bộ định tuyến bị cấm trên đường định tuyến và trạng thái của các bộ định tuyến lân cận để cập nhật thông tin định tuyến phù hợp. Khối ***Update*** sử dụng hai bộ đệm với độ sâu bằng 1 để lưu tạm thời flit đến và flit sau xử lý để bảo đảm tính ổn định và tin cậy của quá trình xử lý và chuyển tiếp thông tin.

3.4 Mô hình hóa, kiểm chứng và thực thi

Hoạt động nghiên cứu thiết kế vi mạch luôn cần hướng đến các sản phẩm hoàn thiện được thực thi trên phiến silicon. Quá trình nghiên cứu này cần thực hiện qua nhiều công đoạn, từ những ý tưởng giải pháp ban đầu thông qua các mô tả kiến trúc và mô phỏng kiểm nghiệm ở nhiều mức độ mới có thể hoàn thiện cho phép thực hiện thực tế. Hình 3.18 mô tả toàn bộ các bước cần thực hiện của giải pháp tái cấu hình mạng trên chip.



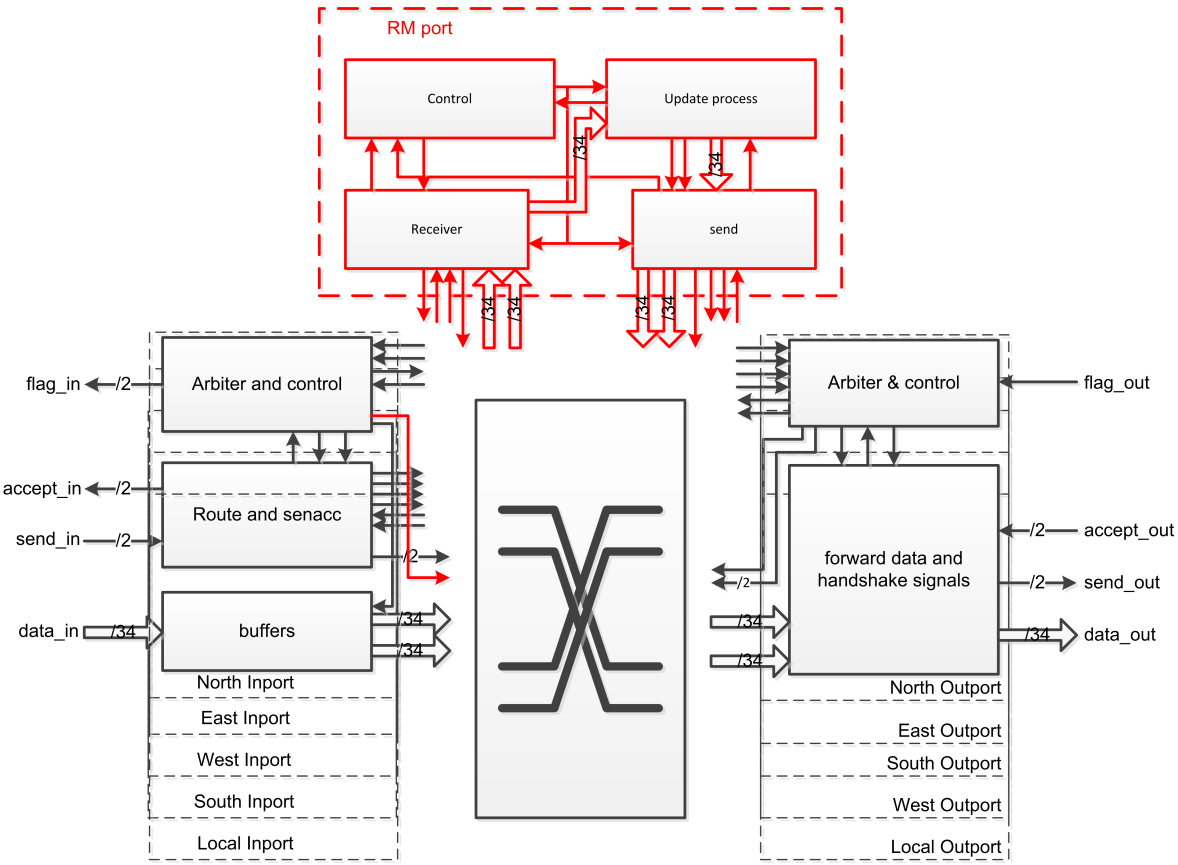
Hình 3.18: Quá trình thực hiện giải pháp mạng trên chip tái cấu hình.

Giải pháp cập nhật thông tin định tuyến sau khi được xây dựng ở mức cao cho phép đánh giá nhanh khả năng thích ứng với các thay đổi cấu hình của mạng trên chip [35]. Kiến trúc mức cao của bộ định tuyến đã thể hiện các chức năng hoạt động và tính khả thi của giải thuật cập nhật định tuyến cho mạng trên chip tái cấu hình. Tuy nhiên, kiến trúc mức cao dựa trên SystemC sẽ rất khó để tổng hợp lên vi mạch và không hiệu quả cho quá trình tối ưu thiết kế. Do vậy, giải pháp cập nhật định tuyến cần được mô tả ở mức thấp hơn theo định hướng tổng hợp trên vi mạch dựa vào các công cụ chuyên dụng cho thiết kế vi mạch. Mục này tập trung trình bày quá trình mô hình hóa kiến trúc bộ định tuyến cho giải pháp tái cấu hình mạng trên chip ở mức chuyển dịch thanh ghi (RTL) bằng cách sử dụng ngôn ngữ mô tả phần cứng VHDL. Sau khi hoàn

thiện quá trình mô phỏng đánh giá kiến trúc bộ định tuyến ở mức RTL, bộ định tuyến được tổng hợp thiết kế bằng công cụ biên dịch vi mạch chuyên dụng (*Design Compiler*) của hãng Synopsys.

3.4.1 Mô hình hóa kiến trúc bộ định tuyến tái cấu hình

Dựa vào các đặc điểm kiến trúc của bộ định tuyến đã được mô tả ở mức cao thông qua các khối cổng vào/ra và khối cổng ảo cập nhật thông tin định tuyến. Trong mục này tập trung thể hiện kiến trúc mức dịch chuyển thanh ghi giữa các khối cổng để thể hiện được đặc điểm thiết kế. Do vậy, bộ định tuyến cho mạng trên chip tái cấu hình vẫn gồm 5 khối cổng vào/ra và khối ***RMport*** nhưng được kết nối qua một ma trận tín hiệu nối chéo. Ma trận tín hiệu nối chéo được xây dựng tách biệt với các khối cổng vào/ra nhằm làm rõ hơn hoạt động dịch chuyển dữ liệu bên trong bộ định tuyến của hoạt động mô tả ở mức RTL. Hình 3.19 mô tả kiến trúc bộ định tuyến cho mạng trên chip tái cấu hình dựa theo nguyên lý luồng tin (*data path*) bằng ngôn ngữ mô tả phần cứng VHDL.



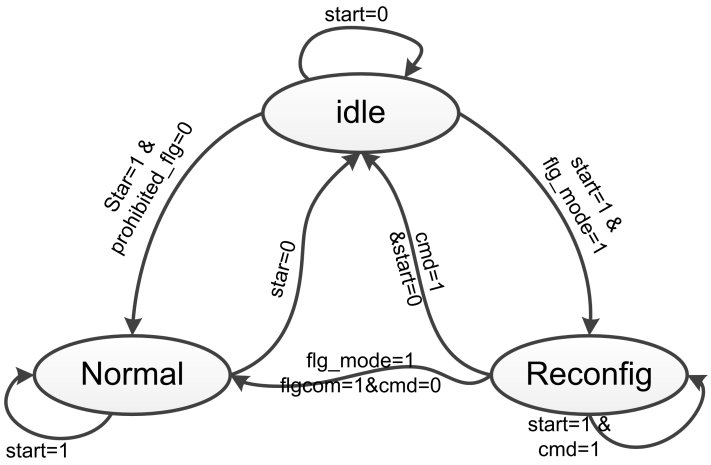
Hình 3.19: Chi tiết kiến trúc bộ định tuyến cho RNoC.

Ở mức RTL, bộ định tuyến tập trung vào hoạt động dịch chuyển dữ liệu giữa các khối chức năng dựa trên nguyên tắc chuyển trạng thái được thể hiện

dưới dạng máy trạng thái hữu hạn. Với mục đích tối ưu hoạt động của bộ định tuyến trong cả hai trường hợp: hoạt động bình thường và hoạt động cập nhật thông tin định tuyến, bộ định tuyến ở mức RTL vẫn hoạt động dựa trên nguyên lý máy trạng thái như trong Hình 3.12. Tuy nhiên, để đảm bảo khả năng đồng bộ giữa các khối chức năng bên trong thiết kế này chỉ cho phép chuyển dịch dữ liệu theo sườn dương của xung nhịp hệ thống.

Khối cổng lối vào - INPORT

Với chức năng nhận tin đầu vào của bộ định tuyến, khối cổng lối vào luôn sẵn sàng nhận tin khi đang ở trạng thái chờ (***idle***) và sẽ chuyển sang một trong hai trạng thái còn lại (***normal*** hoặc ***reconfig mode***) dựa vào quyết định của khối điều khiển như mô tả ở Hình 3.20. Các chế độ hoạt động của khối cổng lối vào được xác định như sau:



Hình 3.20: Sơ đồ trạng thái của máy trạng thái hữu hạn cho khối cổng lối vào.

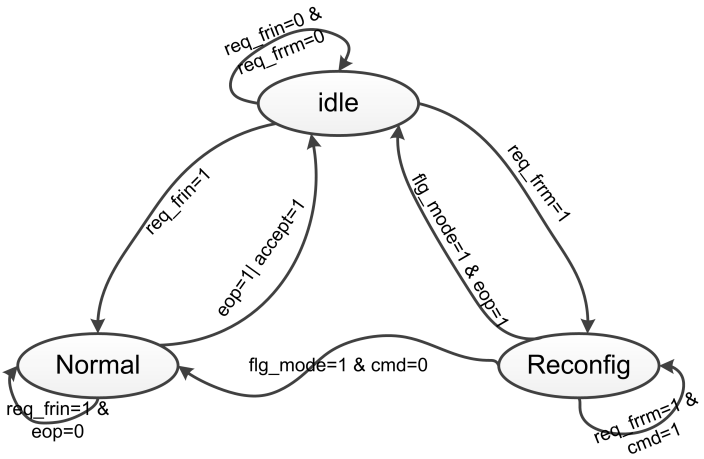
- **Chế độ chờ (*idle*)**: Là trạng thái sẵn sàng của khối cổng lối vào ngay sau khi hệ thống được thiết lập hoạt động, với bộ đệm bên trong ở trạng thái rỗng và tín hiệu điều khiển ở trạng thái mặc định sẵn sàng nhận thông tin đến.

- **Chế độ bình thường (*normal mode*)**: Là trạng thái khối cổng lối vào đang hoạt động chuyển tiếp thông tin từ giao tiếp của bộ định tuyến hướng đến khối cổng lối ra dựa vào thông tin định tuyến đã có.

- **Chế độ tái cấu hình (*reconfig mode*)**: Là trạng thái khối cổng lối vào nhận được flit tiêu đề và thông tin định tuyến trỏ đến lối ra đã bị chặn do bộ định tuyến liền sau bị cấm. Trong chế độ định tuyến, thông tin nhận được ở khối cổng lối vào sẽ được chuyển hướng đến khối cổng ảo cập nhật

RMport, đồng thời dựa vào sự điều khiển của khối cổng ảo ***RMport*** để thực hiện các hoạt động chuyển tiếp thông tin phù hợp.

Khối cổng lối ra - OUTPORT Nhằm tối ưu cho kiến trúc thực thi của bộ định tuyến, khối cổng lối ra được thiết kế tối giản với ba chế độ hoạt động tương ứng với các chế độ hoạt động của bộ định tuyến. Hoạt động của khối cổng lối ra chỉ tập trung chuyển tiếp thông tin khi có yêu cầu chuyển tin từ các khối cổng đầu vào hoặc khối cổng ảo ***RMport*** như mô tả trong Hình 3.21.



Hình 3.21: Sơ đồ trạng thái của máy trạng thái hữu hạn cho khối cổng lối ra.

- **Chế độ chờ (*idle*)**: là chế độ hoạt động của khối cổng lối ra sau khi hệ thống được thiết lập hoạt động nhưng chưa có các yêu cầu chuyển tiếp thông tin từ các khối cổng lối vào hoặc khối cổng ảo ***RMport***. Ở chế độ chờ, khối cổng lối ra luôn sẵn sàng chuyển tiếp thông tin để đáp ứng các yêu cầu truyền tin từ các khối cổng lối vào và khối cổng ảo ***RMport***. Tuy nhiên để bảo đảm hoạt động chuyển tin là tin cậy, khối cổng lối ra chỉ tích cực các tín hiệu bắt tay sau khi nhận được yêu cầu phân xử và xác định được cổng lối vào hoặc ***RMport*** được phép chuyển tiếp thông tin.

- **Chế độ bình thường (*normal mode*)**: là chế độ hoạt động cơ bản của khối cổng lối ra để đáp ứng các yêu cầu chuyển tiếp thông tin từ các khối cổng lối vào của cùng bộ định tuyến đến bộ định tuyến liền sau. Ở chế độ bình thường, khối con ***out_ arb*** thực hiện hoạt động phân xử yêu cầu truyền thông của các khối cổng ảo để đáp ứng truyền thông tối ưu và bảo đảm chất lượng dịch vụ truyền thông cao nhất.

- **Chế độ tái cấu hình (*reconfig mode*)**: là chế độ hoạt động chuyển tiếp thông tin từ khối cổng ảo ***RMport*** do nhận được yêu cầu từ khối cổng

Bảng 3.1: Bảng các trạng thái hoạt động của cổng vào và cổng ra.

Trạng thái	**Mô tả cho cổng vào - IN-PORT**	**Mô tả cho cổng ra - OUT-PORT**
Trạng thái chờ - ***Idle***	Trạng thái sẵn sàng nhận tin của khối cổng lối vào, tín hiệu accept tích cực	Trạng thái rỗi của cổng lối ra để sẵn sàng đáp ứng chuyển tiếp thông tin từ các cổng lối vào hoặc khối cập nhật định tuyến
Chế độ bình thường - ***Normal mode***	Chế độ hoạt động nhận tin từ node mạng phía trước và hướng đến cổng lối ra được lựa chọn	Hoạt động đáp ứng chuyển tiếp thông tin đến node mạng tiếp theo trên đường định tuyến
Chế độ tái cấu hình - ***Reconfig mode***	Khối cổng lối vào nhận được gói tin hướng đến cổng lối ra bị chặn và cần được xử lý cập nhật đường định tuyến phù hợp Trường hợp chỉ xử lý cập nhật flit tiêu đề thì chế độ tái cấu hình sẽ chuyển sang chế độ bình thường của cổng lối vào	Khối cổng lối ra xử lý yêu cầu chuyển tiếp thông tin từ khối cổng ảo và đáp ứng chuyển tiếp thông tin theo điều khiển của khối cập nhật định tuyến Chế độ tái cấu hình sẽ tự chuyển đến chế độ bình thường khi lệnh điều khiển cmd=0 hoặc chuyển tiếp toàn bộ gói tin khi lệnh cmd=1

ảo ***RMport***. Khi có yêu cầu chuyển tiếp thông tin từ khối cổng ảo, khối con điều khiển ***out_ ctrl*** sẽ chờ cho khối cổng lối ra hoàn thành đáp ứng chuyển tiếp gói tin hiện tại để phát lệnh chuyển chế độ hoạt động và điều khiển các khối con bên trong khối cổng vào để đáp ứng yêu cầu chuyển tiếp thông tin của khối cổng ảo ***RMport***.

Khối cổng ảo thực thi tái cấu hình - RMport được xây dựng dựa trên kiến trúc đề xuất mức cao với 4 khối con: *Control, Update, Receiver* và *Send*. Tuy nhiên, kiến trúc chi tiết đã được thực hiện tuân thủ nguyên tắc dịch chuyển thanh ghi dựa trên nguyên lý đồng bộ theo sườn dương của xung nhịp hệ thống. Đồng thời trong thiết kế này cũng đã tách chức năng hợp kênh ở khối ***Receiver*** và phân kênh ở khối ***Send*** vào trong ma trận tín hiệu nối chéo ***Crossbar***.

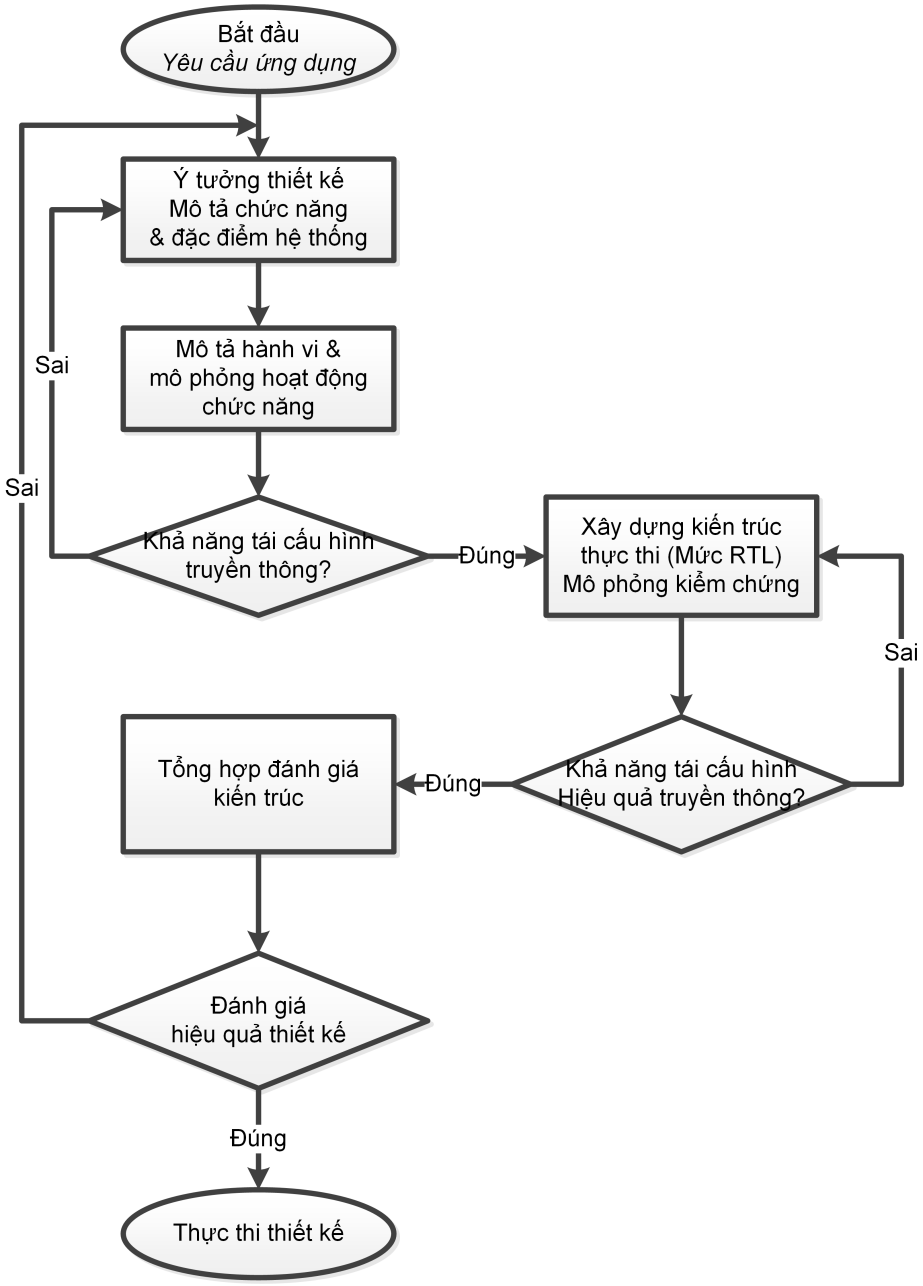
Khối cổng ảo thực thi bộ định tuyến luôn ở chế độ chờ và không hoạt động nếu các cổng lối ra không bị thiết lập trạng thái cấm. Khối cổng ảo này chỉ hoạt động khi có yêu cầu tái cấu hình từ khối cổng lối vào và sẽ trở về trạng thái chờ khi không còn yêu cầu tái cấu hình.

3.4.2 Thực thi kiến trúc

Hoạt động nghiên cứu thiết kế hệ thống cần trải qua một quy trình gồm nhiều giai đoạn để hoàn thành sản phẩm như trong Hình 3.22. Chu trình này

được bắt đầu từ những yêu cầu ứng dụng để hình thành nên các ý tưởng thiết kế, mô tả chức năng và đặc điểm cần có để xây dựng hệ thống. Trên cơ sở phân tích yêu cầu và các chức năng cụ thể, hệ thống cần được mô tả ở mức cao thông qua các hành vi cần có bằng các ngôn ngữ mô tả phần cứng phù hợp nhằm kiểm chứng nhanh các ý tưởng và khả năng triển khai. Quá trình kiểm tra đánh giá cho phép kiểm chứng khả năng đáp ứng yêu cầu hoạt động của thiết kế và cũng là cơ sở để triển khai các hoạt động thiết kế kiến trúc chi tiết của hệ thống. Kiến trúc chi tiết của hệ thống cần được mô tả cụ thể ở mức RTL cho phép thực thi thiết kế cũng cần được đánh giá hoạt động chức năng và khả năng tổng hợp lên vi mạch thông qua các công cụ thiết kế. Ở bước cuối cùng của chu trình là hoạt động tổng hợp thiết kế dựa vào công cụ kết hợp với công nghệ chế tạo phù hợp để tạo nên các vi mạch cụ thể.

Giải pháp cập nhật thông tin định tuyến cho mạng trên chip tái cấu hình cũng được thực hiện đầy đủ các bước tuần tự theo chu trình thiết kế như ở Hình 3.22. Trên cơ sở phân tích hoạt động truyền thông mạng trên chip để hình thành ý tưởng cập nhật thông tin định tuyến cho phép linh hoạt đường đi của thông tin khi có sự thay đổi tái cấu hình. Mạng trên chip tái cấu hình được mô tả hoạt động ở mức cao để khảo sát hoạt động và đánh giá nhanh hiệu quả truyền thông. Từ kết quả đạt được của quá trình đánh giá giải pháp đề xuất, hiệu quả truyền thông của giải pháp là khả quan và đáp ứng được yêu cầu đặt ra cho cả hai trường hợp mạng bình thường và mạng có thay đổi cấu hình. Kiến trúc bộ định tuyến đã được mô tả chi tiết ở mức khả thi và đánh giá hoạt động truyền thông đều đã cho kết quả tương đương với mô hình ở mức cao (Mục 3.4.1). Trong mục này trình bày hoạt động tổng hợp kiến trúc bộ định tuyến đã được mô tả, kết quả tổng hợp bộ định tuyến phản ánh đặc điểm thiết kế và chi phí thực thi và tính khả thi của toàn bộ giải pháp. Ở mức chuyển dịch thanh ghi - RTL, kiến trúc bộ định tuyến được xây dựng một cách chi tiết để thể hiện được các quá trình chuyển tiếp thông tin qua các khối con bên trong bằng các máy trạng thái hữu hạn cũng được tiến hành mô phỏng đánh giá để khẳng định tính khả thi của kiến trúc. Hoạt động tổng hợp kiến trúc đã được xây dựng bằng công cụ chuyên dụng cho phép đánh giá được khả năng thực hiện thiết kế tạo thành vi mạch thực tế đồng thời cũng chỉ rõ hiệu quả thiết kế mang lại thông qua các thông số cụ thể: không gian thực thi (*area cost*), công suất tiêu hao, độ trễ của quá trình chuyển tiếp thông tin qua bộ định tuyến, ... Trên cơ sở kết quả thu được từ công cụ biên dịch thiết kế sẽ cho kết quả là kiến trúc khả thi dựa trên các cổng logic cơ bản; từ kết quả này bộ định tuyến lại tiến hành mô



Hình 3.22: Quy trình nghiên cứu, thiết kế và thực hiện vi mạch số.

Bảng 3.2: Bảng so sánh kết quả thực thi bộ định tuyến của RNoC và một số công trình khác.

STT	**Công trình**	**Công nghệ**	**Không gian thực thi**	**Tần số *MHz***	**Công suất**
1	ReNoC [61, 62]	$90nm$	$580000\mu m^2$	100	$4{,}69mW$
2	BiNoC [9]	$90nm$	$48666gates$	500	$29{,}89mW$
3	BiLink [75]	$65nm$	$219470\mu m^2$	N/A	$28{,}9mW$
4	LBDR [59]	$65nm$	$20355\mu m^2$	1300	N/A
5	DyAD [27]	$160nm$	$25971gates$	333	N/A
6	PRouter[11]	$180nm$	$337471\mu m^2$	100	$9{,}34mW$
7	RRAFT[73]	$90nm$	$680000\mu m^2$	N/A	N/A
8	Kiến trúc đề xuất - RNoC	$130nm$	$461345\mu m^2$ ($54359gates$)	100	$7{,}58mW$

phỏng đánh giá hoạt động để kiểm tra khả năng hoạt động và hiệu quả thiết kế trong điều kiện rất sát thực tế.

Hoạt động tổng hợp kiến trúc đề xuất được thực hiện trên công cụ biên dịch thiết kế (*Design Compiler*) của hãng công nghệ Synopsys, trên nền tảng thư viện CMOS $130nm$ của Global Foundry tại Phòng thí nghiệm trọng điểm Hệ thống tích hợp thông minh (SISLAB). Trên cơ sở tối ưu cho chi phí không gian thiết kế và công suất tiêu thụ, kết quả thu được thông qua báo cáo của phần mềm tổng hợp được trình bày và so sánh với một số công trình tái cấu hình khác ở Bảng 3.2. Do đặc thù của lĩnh vực thiết kế hệ thống, hoạt động thực thi để tạo nên sản phẩm là vi mạch cần một thiết kế hoàn thiện với đầy đủ các chức năng: vi xử lý, bộ nhớ, các lõi IP, kiến trúc truyền thông... Nhưng giải pháp tái cấu hình mạng trên chip đề xuất chỉ tập trung cho hoạt động truyền thông bên trong của hệ thống, mà thiếu đi các thành phần thực hiện chức năng hoạt động để hướng đến ứng dụng cụ thể và tạo nên một vi mạch hoàn chỉnh. Tuy nhiên, kết quả tổng hợp được trình bày ở Bảng 3.2 đã chứng tỏ rằng giải pháp đề xuất là hoàn toàn khả thi để thực hiện chức năng tổng hợp lên vi mạch. Đồng thời, các chi phí không gian thực thi và thông số hoạt động cũng đã cho thấy sự tối ưu của chi phí không gian thực thi so với công trình ReNoC [61, 62], RRAFT [73]. Mặc dù chi phí không gian thực thi của kiến trúc đề xuất - RNoC là lớn hơn so với bộ định tuyến tối ưu cho công suất PRouter [11], nhưng lại có công suất tiêu hao là nhỏ hơn khi chỉ tiêu hao $7{,}58mW$ so với $9{,}34mW$. Hơn nữa, bộ định tuyến PRouter [11] chỉ hướng đến mục tiêu tối ưu công xuất tiêu thụ và chi phí không gian mà không có khả năng thay đổi hoạt động truyền thông để thích ứng với sự thay đổi cấu hình của mạng trên chip. Kết quả tổng hợp của giải pháp đề xuất chưa tối ưu so với các công trình [59, 75] do sử dụng

công nghệ có sự khác biệt lớn, và nhất là đối với công trình [75] tập trung khả năng tái cấu hình ở liên kết giữa hai bộ định tuyến và không bổ sung khả năng tái cấu hình cho bộ định tuyến.

3.5 Kết luận chương

Trên cơ sở nghiên cứu hoạt động truyền thông mạng trên chip và định hướng tái cấu hình trong các thiết kế truyền thông cho các hệ thống trên chip, chương này đã tập trung trình bày giải pháp tái cấu hình mạng trên chip. Hướng trọng tâm đến cấu trúc dạng lưới hai chiều đang rất phổ biến của nhiều nghiên cứu mạng trên chip, giải thuật cập nhật định tuyến được đề xuất cho phép mạng trên chip có khả năng tự thích ứng với sự thay đổi cấu hình do có bộ định tuyến rời khỏi mạng. Việc lựa chọn hoạt động cập nhật đường định tuyến để hỗ trợ cho giải thuật định tuyến tĩnh tại nguồn của các mạng trên chip cho phép tận dụng được tính ổn định và hiệu quả truyền thông của giải thuật định tuyến tĩnh nhưng lại đáp ứng linh hoạt cho những thay đổi trong các trường hợp tái cấu hình. Hơn nữa, giải thuật cập nhật định tuyến chia nhỏ các trường hợp cập nhật định tuyến góp phần tối ưu trong quá trình xử lý thay đổi đường định tuyến và phù hợp với các điều kiện ràng buộc của việc thực thi trên vi mạch.

Đáp ứng mục tiêu thực hiện giải pháp tái cấu hình mạng trên chip, kiến trúc bộ định tuyến đã được đề xuất ở mức cao cho phép triển khai nhanh quá trình đánh giá hiệu quả của giải thuật cập nhật định tuyến. Mô tả kiến trúc ở mức cao bằng ngôn ngữ mô tả *SystemC* thể hiện đầy đủ các thành phần chức năng cũng như hoạt động bên trong của bộ định tuyến thông qua quá trình mô phỏng đánh giá. Điều này góp phần quan trọng để khẳng định tính khả thi của giải pháp tái cấu hình mạng trên chip với khả năng thay đổi đường định tuyến linh hoạt thích ứng với nhiều trường hợp có bộ định tuyến rời khỏi mạng. Giải pháp tái cấu hình bằng cách cập nhật thông tin định tuyến đã được công bố tại công trình [J1] trên tạp chí chuyên ngành JEC của Hội Vô tuyến-Điện tử Việt Nam năm 2014 [J1].

Hướng đến mục tiêu hoàn thiện giải pháp tái cấu hình cho các hệ thống phức hợp, chương này đã trình bày theo một quy trình nghiên cứu thiết kế mạng trên chip tái cấu hình như trong Hình 3.18. Trong đó quá trình mô tả chi tiết thiết kế ở mức dịch chuyển thanh ghi - RTL bằng ngôn ngữ mô tả phần cứng - VHDL và thể hiện nguyên lý hoạt động cụ thể của các khối cổng bên trong bộ định tuyến bằng các máy trạng thái hữu hạn - FSM. Việc tổng hợp kiến trúc bộ

định tuyến là bước đi quan trọng để chứng minh hiệu quả và tính khả thi của giải pháp tái cấu hình đã đưa ra. Bằng cách sử dụng công cụ tổng hợp chuyên dụng là phần mềm *Design Complier* của Synopsys trên nền công nghệ chế tạo CMOS 130*nm* của hãng Global Foundry, các tham số kết quả của thiết kế cho thấy rõ hiệu quả kiến trúc bộ định tuyến như đã trình bày trong Bảng 3.2.

Chương 4

Đánh giá hiệu năng truyền thông và hiệu quả giải pháp tái cấu hình

Như đã đề cập trong các chương trước, mạng trên chip được xem là giải pháp truyền thông hiệu quả, hứa hẹn thay thế các giải pháp truyền thông truyền thống trong các hệ thống phức hợp, có nhu cầu truyền thông lớn. Tuy vậy, hiệu năng hoạt động của mạng trên chip lại phụ thuộc rất nhiều vào các thông số của mạng. Đồng thời, việc mô phỏng và đánh giá hiệu năng truyền thông của mạng trên chip cần được quan tâm phát triển trước khi thực thi với các ứng dụng cụ thể. Hoạt động mô phỏng đánh giá cho phép người thiết kế nắm bắt được các yếu tố của mạng trên chip như (công suất tiêu thụ, chi phí không gian thực thi,...) để tạo nên kiến trúc mạng trên chip thực sự phù hợp với ứng dụng. Hiện nay, các công cụ mô phỏng và đánh giá đã và đang được phát triển hướng đến nhiều mục tiêu để phục vụ cho những định hướng nghiên cứu cụ thể. Tuy nhiên, các công cụ đã có được phát triển cho từng mục tiêu cụ thể và rất khó tùy biến cho các đánh giá cụ thể người sử dụng. Vậy nên, sau khi tìm hiểu về hoạt động đánh giá và các công cụ sẵn có, việc phát triển một công cụ riêng phù hợp với định hướng nghiên cứu mạng trên chip và định hướng tái cấu hình. Công cụ đánh giá phải linh hoạt trong đánh giá hoạt động truyền thông cho mạng trên chip trước và sau khi sử dụng giải pháp tái cấu hình. Đồng thời công cụ phải tạo được nhiều kịch bản đánh giá phù hợp vụ mục tiêu khảo sát hoạt động truyền thông mạng trên chip và mạng trên chip tái cấu hình.

Chương này sẽ trình bày ý tưởng và phương pháp xây dựng hệ thống mô phỏng, đánh giá hiệu năng mạng trên chip dựa vào SystemC – một ngôn ngữ mô tả phần cứng bậc cao – cho phép đánh giá nhanh hiệu năng của mạng trên chip đề xuất. Việc đánh giá mạng trên chip thông thường và mạng trên chip có khả năng tái cấu hình sẽ lần lượt được đề cập trong các phần tiếp theo của chương. Các kết quả đánh giá hiệu năng sẽ được so sánh với nhau và với các kết quả của

các nghiên cứu trước đây nhằm làm rõ hiệu quả của mô hình mạng trên chip có khả năng tái cấu hình đề xuất.

4.1 Hoạt động đánh giá hiệu năng truyền thông mạng trên chip

Để biết được một mô hình mạng trên chip đề xuất có thực sự đem lại hiệu quả cao hay không thì việc nghiên cứu đề xuất các phương pháp đánh giá hiệu năng của mạng là thực sự cần thiết. Đây cũng là một chủ đề được nhiều nhóm nghiên cứu quan tâm. Trong phần này chúng ta sẽ đề cập đến thông số phản ánh hiệu năng truyền thông của một mô hình mạng và các phương pháp đánh giá các thông số đó.

4.1.1 Thông số đánh giá hiệu năng truyền thông

Chức năng chính của mạng trên chip là đảm bảo khả năng truyền thông phù hợp với mọi hoạt động trao đổi thông tin giữa các lõi chức năng. Và hiện nay, mạng trên chip đang được tập trung nghiên cứu với nhiều định hướng kiến trúc và lĩnh vực ứng dụng. Nghiên cứu hoạt động đánh giá hiệu năng truyền thông cho phép đánh giá hiệu quả hoạt động truyền thông của mạng trên, đồng thời quá trình này cũng làm cơ cho việc khảo sát hoạt động truyền thông của mạng trên chip để hướng đến giải pháp cho mạng trên chip tái cấu hình.

Trong nhiều thông số đánh giá hoạt động truyền thông qua mạng thì **độ trễ truyền** (*Latency*) được quan tâm và áp dụng trong hầu hết các khảo sát đánh giá truyền thông nói chung và lĩnh vực mạng trên chip càng được chú trọng hơn. Xét trong mạng truyền gói, độ trễ truyền là giá trị thời gian lan truyền của một giao tác truyền thông. Trong hoạt động truyền thông của mạng trên chip, gói tin được chia nhỏ thành các đơn vị thông tin gọi là flit nên trễ truyền lúc này được tính bằng khoảng thời gian bắt đầu từ nguồn phát đưa flit tiêu đề vào mạng đến lúc đích đến nhận gói tin hoàn thiện như trong biểu thức (4.1) [53].

$$L_i = time_{receiver} - time_{send} \tag{4.1}$$

Trong đó: $time_{receiver}$ là thời điểm khối đích nhận đầy đủ gói tin.

$time_{send}$ là thời điểm khối nguồn bắt đầu phát tin vào mạng.

Và áp dụng cho mạng trên chip với yêu cầu đánh giá tổng thể trên toàn mạng để chỉ ra khả năng đáp ứng cho hoạt động của hệ thống; giá trị độ trễ

truyền toàn mạng là giá trị trung bình cho tất cả các giao tác truyền thông được thực hiện như trong biểu thức (4.2) [53]:

$$L_{avg} = \frac{\sum_{i=1}^{P} L_i}{P} \tag{4.2}$$

Trong đó, P là tổng số gói tin được gửi trong quá trình đánh giá hoạt động truyền thông của mạng.

Với đặc điểm hoạt động theo xung nhịp của các hệ thống trên chip, giá trị của độ trễ truyền có thể tính theo chu kỳ xung nhịp (*clk*). Trong một số công trình thực hiện ở mức tổng hợp vi mạch thì giá trị thời gian sẽ được tính theo thời gian thực truyền của các giao tác truyền thông.

Trong truyền thông trên chip hiện tại, việc đánh giá khả năng truyền thông thường thể hiện qua giá trị băng thông (*bandwidth*) là số lượng thông tin truyền qua trong một đơn vị thời gian. Tuy nhiên giá trị băng thông này thường chỉ dùng để chỉ khả năng truyền tải qua một giao diện liên kết, như với BUS chia sẽ, hoặc liên kết điểm-điểm. Trong mạng trên chip, thông tin trao đổi qua nhiều liên kết đồng thời, nên để xác định hiệu quả truyền thông cần thể hiện cho tất cả các liên kết mạng. Do đó, trong mạng trên chip sử dụng giá trị trung bình của lượng tin truyền qua liên kết mạng trong một đơn vị thời gian; đại lượng đó được gọi là **thông lượng mạng (*throughput*)**. Từ đó cho thấy giá trị thông lượng mạng được dùng để đánh giá khả năng đáp ứng truyền thông của mạng trên chip. Thông lượng mạng được xác định theo như biểu thức dưới đây [53]:

$$TP = \frac{(TotalPackets) \times (PacketSize)}{(NumberofIPcore) \times (TotalTime)} \tag{4.3}$$

Trong đó, *TotalPackets* là tổng số lượng gói tin được gửi vào mạng.

PacketSize là kích thước gói tin tính theo đơn vị cơ sở truyền thông qua mạng, được tính là số flit trên mỗi gói tin.

NumberofIPcore là số lượng các lõi IP của hệ thống phát/nhận tin.

TotalTime là tổng thời gian để mạng truyền hết tất cả các gói tin.

Đơn vị của thông lượng trong biểu thức (4.3) được tính bằng lượng đơn vị tin qua giao diện trong một đơn vị thời gian. Do vậy trong trường hợp mô phỏng hoạt động dựa theo xung nhịp của mạng trên chip thì đơn vị thông lượng sẽ là *flit/clk*. Trường hợp mạng trên chip được đánh giá và tổng hợp thành mạch điện, hoạt động truyền thông tính theo đơn vị thời gian thực thì đơn vị tính của thông lượng cũng được tính theo số lượng *flit/đơn vị thời gian*.

4.1.2 Phương pháp đánh giá

Hoạt động đánh giá truyền thông là một quy trình bắt buộc để khẳng định thiết kế đáp ứng tốt các đặc điểm trao đổi thông tin giữa các thành phần của hệ thống. Do đó, mỗi hệ thống với những định hướng hoạt động khác nhau sẽ có nhiều yêu cầu truyền thông rất đa dạng. Trong lĩnh vực thiết kế trên chip với các hệ thống phức hợp, đa dạng thì vấn đề truyền thông trên chip càng phức tạp hơn với nhiều yêu cầu về trao đổi dữ liệu song song thời gian thực, khối dữ liệu lớn, có nhiều mức ưu tiên ... Vậy nên, hoạt động đánh giá mạng cần có những chiến lược đánh giá phù hợp với mục tiêu nghiên cứu đồng thời bảo đảm được độ tin cậy của hoạt động đánh giá. Chiến lược đánh giá là tổng hợp các quy luật tạo thông tin trao đổi, điều khiển luồng tải tin vào mạng và sử dụng các tham số đánh giá phù hợp với mục tiêu đánh giá đề ra. Do vậy, để hướng đến các ứng dụng của hệ thống thực, các đánh giá mạng trên chip cần những chiến lược đánh giá phù hợp bằng việc sử dụng một mô hình tải tin kết hợp với các tham số hiệu năng tương ứng. Trong hệ thống trên chip với nhiều lõi IP đa dạng (vi xử lý, bộ nhớ, DSP ...), mỗi lõi IP sẽ có những yêu cầu trao đổi thông tin khác nhau và thay đổi theo thời gian. Do đó, trong mô phỏng đánh giá hiệu năng truyền thông của mạng trên chip cần xác định **mô hình tải tin (*workload model*)** là tập hợp các quy luật tạo tải tin và phân bố các hoạt động trao đổi tin trong hệ thống. Việc mô tả chi tiết mô hình tải tin sẽ được thể hiện thông qua ba tham số là: phân bố cặp nguồn đích, tốc độ phát tải và kích thước gói tin [30].

Phân bố cặp nguồn-đích là quy luật để xác định hoạt động trao đổi thông tin giữa các lõi IP bên trong hệ thống trong quá trình đánh giá. Trong hệ thống hoạt động trao đổi thông tin giữa các lõi IP thường được xác lập theo chức năng cụ thể và cũng mang tính xác suất ngẫu nhiên. Do đó, việc lựa chọn phân bố giữa các lõi IP có tính quyết định đến sự khách quan, độ tin cậy kết quả của hoạt động đánh giá. Hiện nay hoạt động đánh giá thường sử dụng các phân bố cặp nguồn đích như sau:

- **Đảo bit (*bit reversal*)** mỗi lõi IP nguồn với địa chỉ biểu diễn theo chuỗi số nhị phân $a_{n-1}a_{n-2}...a_1a_0$ sẽ phát tin đến lõi IP đích có địa chỉ nhị phân là $a_0a_1...a_{n-2}a_{n-1}$.

- **Dịch bit (*perfect shuffle*)** mỗi lõi IP nguồn với địa chỉ theo chuỗi nhị phân $a_{n-1}a_{n-2}...a_1a_0$ sẽ phát tin đến các lõi IP đích có địa chỉ nhị phân $a_{n-2}a_{n-3}...a_1a_0a_{n-1}$ (dịch vòng trái một *bit*).

- **Chéo cánh bướm (*butterfly*)** là dạng trao đổi thông tin giữa hai lõi IP chéo cánh bằng việc đảo các bit lần lượt. Như một lõi IP có địa chỉ $a_{n-1}a_{n-2}...a_1a_0$ sẽ trao đổi thông tin với lõi IP có địa chỉ $a_0a_{n-2}...a_1a_{n-1}$.

- **Ma trận chuyển vị (*matrix transpose*)** khi đánh giá sử dụng phân bố ma trận chuyển vị thì một lõi IP có địa chỉ nhị phân là $a_{n-1}a_{n-2}...a_1a_0$ sẽ trao đổi thông tin với một lõi IP khác có địa chỉ $a_{\frac{n}{2}-1}...a_0a_{n-1}...a_{\frac{n}{2}}$.

- **Lấy bù (*complement*)** mỗi lõi IP nguồn phát tin đến lõi IP đích có địa chỉ là giá trị bù của địa chỉ đích. Ví dụ: một lõi IP có địa chỉ nhị phân là $a_{n-1}a_{n-2}...a_1a_0$ sẽ phát đến lõi IP có địa chỉ nhị phân là $\overline{a_{n-1}a_{n-2}...a_1a_0}$.

Tốc độ phát tải (*workload*) thể hiện số lượng đơn vị thông tin được phát vào mạng trong mỗi đơn vị thời gian của nguồn tin. Trong đặc điểm mô phỏng đánh giá mạng trên chip hoạt động theo nguyên lý đồng bộ, thời điểm nguồn tin phát vào mạng phụ thuộc vào xung nhịp của hệ thống. Lúc đó tải tin được định nghĩa là tỷ lệ của số lần phát tin trên mỗi chu kỳ xung nhịp (*clk*); và do đó ta có thể sử dụng đơn vị (%) cho tải tin.

Kích thước gói tin (*packet size*) là tham số dùng để chỉ độ dài của gói tin lan truyền trong mạng. Với mạng trên chip sử dụng đơn vị tin cơ bản là **flit** thì kích thước gói tin được tính là số lượng *flit* trong mỗi gói tin. Với các hệ thống trên chip có sự trao đổi phức tạp giữa các thành phần thì kích thước gói tin lan truyền trong mạng sẽ rất đa dạng. Do đó hoạt động mô phỏng đánh giá truyền thông của mạng trên chip cần có chiến lược linh hoạt thông số kích thước gói tin để kết quả đánh giá sát với yêu cầu truyền thông của hệ thống thực.

Trong định hướng nghiên cứu hoạt động đánh giá mạng trên chip với nhiều giải pháp tiếp cận đa dạng, từ các nghiên cứu đánh giá trong định hướng ứng dụng cụ thể [63] đến các công trình xây dựng công cụ chuyên dụng [40]. Hoạt động đánh giá luôn được xem xét dựa trên các mô hình tải tin phù hợp, và đồng thời cũng hướng đến khả năng đánh giá khách quan theo các thông số hiệu năng. Để nghiên cứu đánh giá toàn diện hơn thì Rahmani và cộng sự cũng đã tập trung nghiên cứu các mô hình tải tin hướng đến khả năng phân tích hiệu năng hoạt động của mạng trên chip [56]. Những điều đó cho thấy được tầm quan trọng của chiến lược đánh giá trong các nghiên cứu về hoạt động truyền thông của mạng trên chip.

4.2 Đánh giá hoạt động truyền thông trên chip

Nắm bắt được tầm quan trọng của hoạt động đánh giá cho các kiến trúc truyền thông mạng trên chip, luận án đã tiến hành xây dựng một nền tảng cho hoạt động đánh giá. Hướng đến mục tiêu xây dựng một nền tảng hữu ích và hỗ trợ tối đa cho hoạt động đánh giá khách quan khả năng truyền thông của mạng trên chip, nền tảng đã sử dụng ngôn ngữ mô tả phần cứng mức cao (TLM: *Transaction-Level Modeling*) là *SystemC*. Với đặc điểm mô tả mức cao và khả năng kết hợp với ngôn ngữ C, *SystemC* cho phép xây dựng nhanh các kiến trúc chức năng bên trong của mạng trên chip và nhất là các lõi IP cho mục tiêu đánh giá. Do đó, nền tảng đánh giá được phát triển vừa cho phép đánh giá nhanh khả năng truyền thông của các kiến trúc mạng trên chip vừa có tính năng hỗ trợ đa dạng nhiều kịch bản đánh giá khách quan.

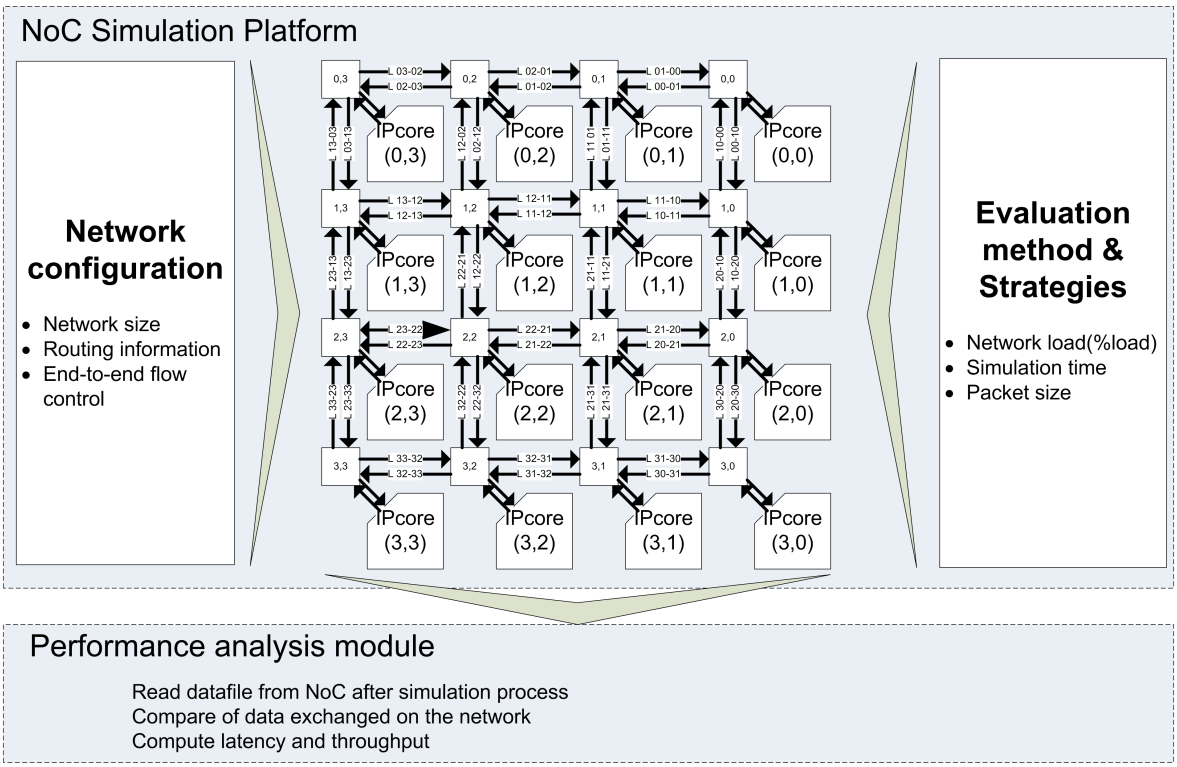
4.2.1 Thiết kế cho đánh giá truyền thông

Nền tảng đánh giá mạng trên chip được xây dựng cho mục tiêu đánh giá hoạt động truyền thông của mạng trên chip với ba thành phần chính được mô tả trong Hình 4.1. Thành phần chính là mạng trên chip được ghép với các lõi IP cho mô phỏng đánh giá được mô tả bằng ngôn ngữ mô tả phần cứng mức cao *SystemC*. Chức năng đánh giá cũng được thiết kế bằng SystemC cho phép mô phỏng đánh giá gắn liền với mạng trên chip. Và trong thiết kế đánh giá này đã sử dụng ngôn ngữ câu lệnh của Linux (*script*) để thực hiện mô phỏng đánh giá hoạt động của mạng trên chip một cách tự động trên nền tảng hệ điều hành Linux [34].

Trong đó mạng trên chip sử dụng cấu trúc liên kết dạng lưới 2D-mesh với kích thước $m \times n$ (trong Hình 4.1 thì $m = n = 4$). Hoạt động trao đổi dữ liệu giữa các bộ định tuyến theo nguyên lý đồng bộ với hai kênh ảo có độ rộng mỗi flit là $34bit$ cho mỗi liên kết một chiều. Với đặc điểm thiết kế cho mô phỏng đánh giá, nền tảng này không sử dụng giao tiếp mạng mà đề xuất kiến trúc lõi IP cho mục tiêu đánh giá phát trực tiếp gói tin có cấu trúc phù hợp với hoạt động truyền thông của mạng trên chip.

Hoạt động mô phỏng quá trình truyền thông của mạng trên chip và đánh giá được điều khiển bởi các chiến lược đánh giá với khả năng tùy biến các thông số bởi ngôn ngữ câu lệnh của Linux.

Với mục tiêu xây dựng nền tảng đáng giá một cách khách quan và toàn diện cho mạng trên chip, khối đánh giá cũng được xây dựng bằng ngôn ngữ



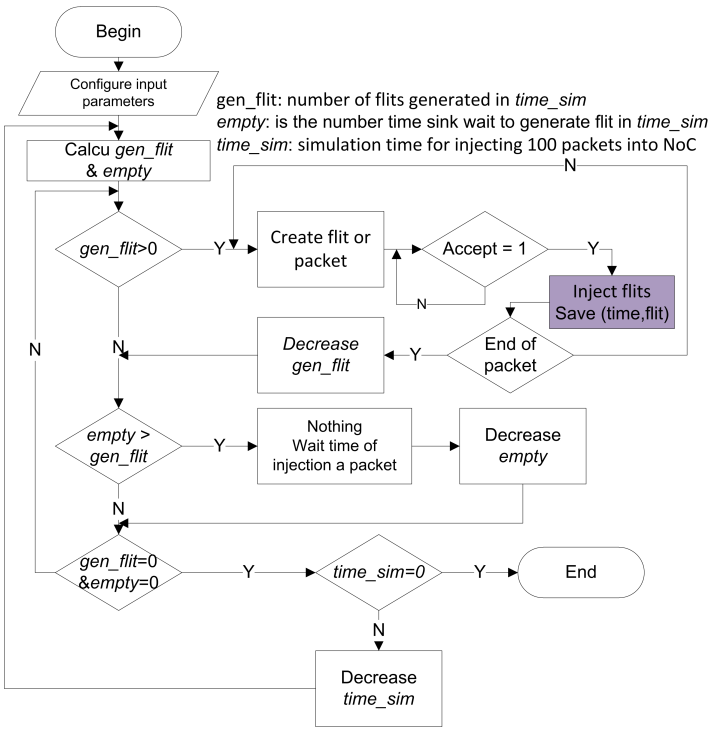
Hình 4.1: Nền tảng đánh giá truyền thông mạng trên chip.

SystemC. Khối đánh giá sẽ phân tích dữ liệu có được của quá trình mô phỏng mạng trên chip được lưu bởi các lõi IP để tính toán giá trị của các đại lượng dùng cho hoạt động đánh giá hiệu năng truyền thông. Do vậy, khối phân tích đánh giá có chức năng đọc các *file* dữ liệu để phân tích và so sánh thông tin trao đổi qua mạng. Kết quả phân tích sẽ cho ra độ tin cậy của thông tin trao đổi, các giá trị về độ trễ truyền và thông lượng mạng.

Hoạt động đánh giá truyền thông mạng trên chip cần thể hiện rõ khả năng truyền thông của thiết kế trong nhiều kịch bản khác nhau. Nhưng trong quá trình thay đổi hệ số tải tin phát vào mạng đã cho thấy có xảy ra tranh chấp tài nguyên cục bộ. Điều này tạo nên sự mất cân đối trong hoạt động phân xử để cung cấp tài nguyên cho các giao tác truyền thông, tạo nên sự mất cân bằng và tác động đến kết quả đánh giá. Để cân bằng giữa yêu cầu tải tin phát vào mạng và khoảng chờ trong một chu trình phát, nền tảng được bổ sung một lõi IP cho mô phỏng đánh giá có khả năng cân bằng thời gian. Hình 4.2 thể hiện thuật toán cân bằng đã được sử dụng tại nguồn phát tin.

4.2.2 Kịch bản đánh giá

Xây dựng nền tảng đánh giá cho mạng trên chip hướng đến một giải pháp đánh giá toàn diện và hiệu quả cho các kiến trúc truyền thông mạng trên chip



Hình 4.2: Lược đồ hoạt động phát tải vào mạng.

với khả năng đánh giá toàn diện và khách quan. Trong điệu kiện mô phỏng đánh giá hoạt động truyền thông, các giải pháp đánh giá cần những cơ chế tạo tải tin đánh giá gần sát với với những đặc điểm hoạt động của hệ thống trong thực tế, cũng như việc tạo các ngữ cảnh đánh giá khả năng tới hạn của kiến trúc truyền thông để cho phép khảo sát khả năng đáp ứng trong nhiều trường hợp khác nhau. Nhằm đáp ứng yêu cầu khảo sát khả năng đáp ứng linh hoạt cho các hoạt động truyền thông phức tạp bên trong các hệ thống phức hợp, nền tảng đánh giá hỗ trợ đa dạng các hoạt động mô phỏng đánh giá thông qua các tham số cấu hình.

Kịch bản đánh giá 1: Cho mỗi kiến trúc mạng cần đánh giá cần được thiết lập cố định kích thước gói tin và tăng dần tốc độ phát tin (*data load*) ở mỗi lõi IP đánh giá; quá trình này sẽ được lặp lại tuần tự cho nhiều kích thước gói tin khác nhau.

Kịch bản đánh giá 2: Tương ứng cho mỗi kiến trúc mạng cần đánh giá, quá trình mô phỏng hoạt động truyền thông được thực hiện lần lượt với cùng tốc độ phát tin nhưng tăng dần kích thước gói tin. Kịch bản này sẽ cho phép chúng ta đánh giá khả năng đáp ứng truyền thông của mạng trên chip đối với mỗi dạng gói tin có kích thước khác nhau thông qua hai thông số chính là: độ trễ truyền và thông lượng mạng.

Trên môi trường của hệ điều hành *Linux* và sử dụng ngôn ngữ dòng lệnh

Bảng 4.1: Thông số của mạng trên chip trong mô phỏng đánh giá

Tham số	**Giá trị**
Các tham số cấu hình và mạng	
Cấu trúc liên kết	2D mesh, kích thước 4×4
Điều kiển luồng	Cơ chế Credit-based
Điều chuyển dữ liệu	Cơ chế Wormhole
Định tuyến	Giải thuật định tuyến XY tại nguồn
Các tham số mô phỏng	
Mô hình tải tin	Phân bố lấy bù, tải đồng dạng
Kích thước gói tin	Từ $1 \div 256$ flits
Tỷ lệ tải tin	$1\% \div 100\%$

(*script*) nền tảng đánh giá hỗ trợ cấu hình phát tải tin tại các lõi IP cho đánh giá theo các kịch bản được sử dụng. Bảng 4.1 liệt kê các thông số cho cấu hình mạng trên chip và hoạt động mô phỏng truyền thông đã được dùng trong việc đánh giá của luận án.

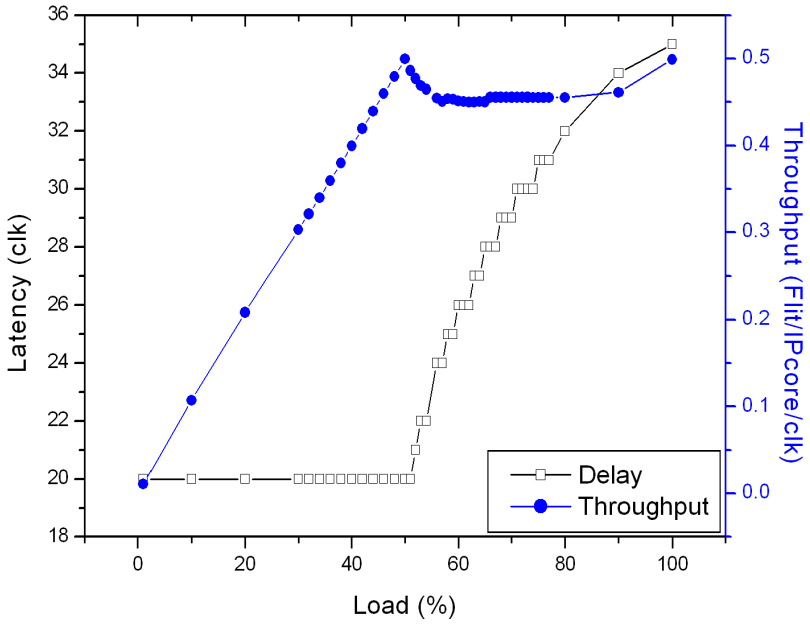
4.2.3 Kết quả đánh giá

Sử dụng nền tảng đánh giá đã được mô tả ở trên (Mục 4.2.1) mạng trên chip được mô phỏng đánh giá dựa theo các kịch bản khác nhau đã thu được các kết quả là độ trễ truyền và thông lượng mạng theo tỷ lệ tải tin.

Hình 4.3 là kết quả đánh giá hiệu năng hoạt động truyền của mạng trên chip với kích thước gói tin là 16*flit* theo độ trễ truyền và thông lượng mạng. Từ kết quả thu được cho thấy độ trễ truyền của mạng ổn định khi tải mạng nhỏ hơn 50% và tăng nhanh khi tải tin phát vào mạng lớn hơn 50%. Độ trễ truyền nhỏ nhất là 20 chu kỳ xung nhịp (*clk*) phù hợp với hoạt động đồng bộ phát từng *flit* và mạng theo xung nhịp của gói tin 16*flit* và thời gian chuyển qua các *hop* trung gian. Với đặc điểm của giải thuật định tuyến tĩnh XY được sử dụng, mạng vẫn hoạt động ổn định khi tải tin tiến đến 100%; và lúc đó độ trễ truyền cực đại là 35 chu kỳ xung nhịp.

Hình 4.3 cũng cho thấy đáp ứng của mạng ổn định khi tải tin dưới 50% phù hợp với giá trị thông lượng mạng tăng tuyến tính theo giá trị tải tin đầu vào. Khi tải tin bắt đầu lớn hơn 50% do sự tranh chấp tài nguyên tăng nhanh nên thông lượng có giảm nhưng không đáng kể, tiến đến giá trị bão hòa khi tải tin tiến đến 100% với giá trị bão hòa là 0,499*flit/IPcore/clk* [36].

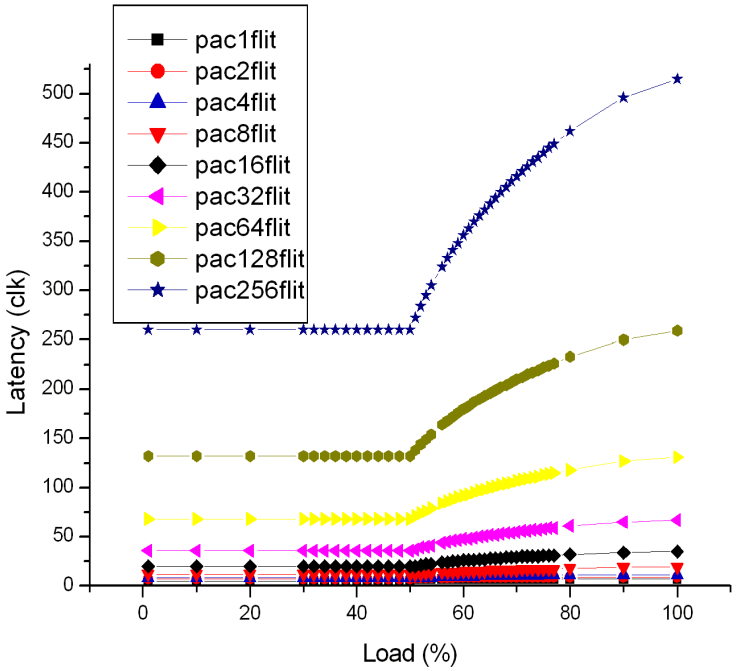
Để đánh giá toàn diện khả năng truyền thông của mạng trên chip hoạt động đánh giá thay đổi chiến lược đánh giá bằng cách tăng dần kích thước gói tin để khảo sát sự phụ thuộc của độ trễ truyền theo tải tin và kích thước gói



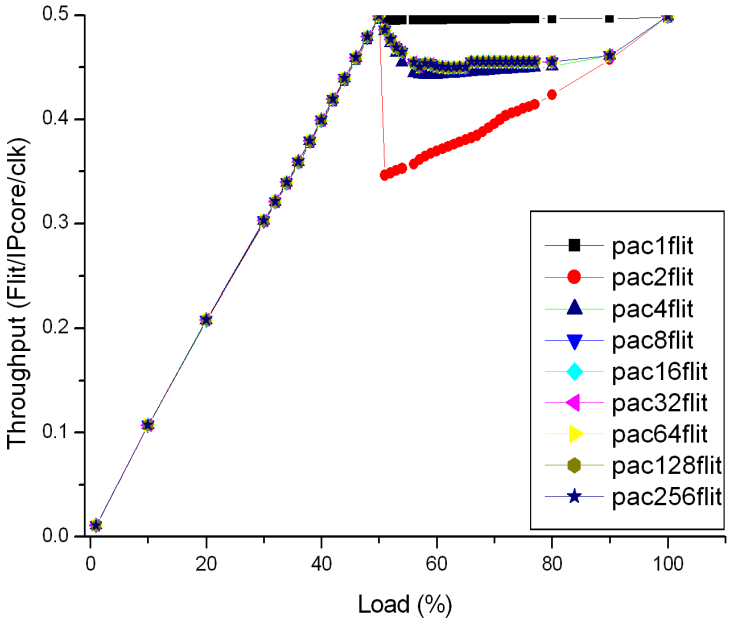
Hình 4.3: Độ trễ truyền và thông lượng theo hệ số tải với gói dữ liệu 16*flit*.

tin. Kết quả của việc đánh giá độ trễ truyền theo kích thước gói tin được biểu diễn ở Hình 4.4. Hình này cho thấy quy luật thay đổi của độ trễ truyền theo tải tin là tương đồng với tất cả các kích thước gói tin khác nhau $(1 \div 256)$. Khi tải tin dưới 50% độ trễ truyền ít thay đổi và có giá trị xấp xỉ với kích thước gói tin và số hop trung bình của giải thuật định tuyến. Khi tải tin tăng lên lớn hơn 50% độ trễ truyền tăng nhanh do đáp ứng của mạng đã bão hòa khi giá trị tải tin tiến đến 100%.

Để khảo sát sự đáp ứng của mạng với các yêu cầu truyền thông, thông lượng mạng cần được đánh giá theo sử thay đổi kích thước gói tin và tải tin đầu vào. Kết quả đánh giá thông lượng mạng khi thay đổi kích thước gói tin được trình bày như trong Hình 4.5. Kết quả nay cho thấy thông lượng mạng thiết kế đáp ứng yêu cầu truyền thông tương tự nhau với tất cả các dạng gói tin có kích thước thay đổi. Khi tải tin có tỷ lệ dưới 50%, thông lượng mạng tăng tuyến tính cùng quy luật đối với tất cả các dạng gói tin. Ở miền bão hòa, giá trị thông lượng mạng có sai khác nhau là do sự phân xử tranh chấp và giải phóng tài nguyên truyền thông. Trong đó gói tin có kích thước 1*flit* ít bị tác động nhất nên có giá trị miền bão hòa không đổi. Các gói tin có kích thước lớn hơn 1*flit* cần có thời gian giải phóng và cấp lại tài nguyên có tỷ lệ với kích thước gói khác nhau nên giá trị miền bão hòa cũng có sự sai khác. Với đặc điểm của giải thuật định tuyến tĩnh XY và cơ chế điều chuyển dữ liệu WH áp dụng cho cấu trúc



Hình 4.4: Trễ truyền thông đánh giá theo hệ số tải và kích thước gói tin.



Hình 4.5: Thông lượng mạng theo hệ số tải và kích thước gói tin.

mạng dạng lưới hai chiều 2D-mesh, trung bình mỗi liên kết mạng sẽ có hai giao tác truyền tin tranh chấp để truyền tin. Do đó, hoạt động phân xử tại các bộ định tuyến luôn có sự tranh chấp tài nguyên và tối đa chỉ có thể đáp ứng được 50% các yêu cầu nếu các giao tác truyền thông liên tục phát tin. Hình 4.5 cho thấy trường hợp gói tin có số *flit* bằng 2 giảm nhanh khoảng 20% và các giá trị kích thước gói khác giảm khoảng 10% khi tải tin lớn hơn 50%. Do vậy, kết quả đánh giá thông lượng thể hiện trong Hình 4.5 cho thấy kiến trúc được đánh giá có thông lượng lớn nhất tiệm cận với giá trị ngưỡng 0,5(*flit/IPcore/clk*) khi tất cả các lõi IP đều yêu cầu truyền tin [48]. Đồng thời, các kết quả này cũng là một minh chứng cho thấy nền tảng đánh giá đáp ứng tốt yêu cầu mô phỏng đánh giá nhanh và khách quan đối với các hoạt động mô phỏng truyền thông và đánh giá hiệu năng của mạng trên chip.

4.3 Đánh giá giải pháp tái cấu hình mạng trên chip

Từ các kết quả thu được của hoạt động đánh giá mạng trên chip đã xây dựng và công bố tại các công trình [C2,C3] như đã trình bày trong Mục 4.2. Nền tảng đánh giá này được tái sử dụng cho cho hoạt động truyền thông của mạng trên chip tái cấu hình với bộ định tuyến đã được trình bày trong Chương 3. Do vậy, trong mục này sẽ tập trung trình bày hoạt động đánh giá áp dụng cho mạng trên chip trên cơ sở những kịch bản đánh giá đã có kết hợp với khả năng mô phỏng các trường hợp tái cấu hình khác nhau.

4.3.1 Phương pháp đánh giá giải pháp tái cấu hình mạng trên chip

Hoạt động truyền thông tái cấu hình mạng trên chip cần được khảo sát đánh giá một cách toàn diện để minh chứng được khả năng đáp ứng linh hoạt các yêu cầu truyền thông của hệ thống trong nhiều điều kiện thay đổi khác nhau. Trên cơ sở các giải pháp đánh giá mạng trên chip được trình bày ở Mục 4.2.2 cho phép linh hoạt thiết lập các thông số trong quá trình mô phỏng hoạt động truyền thông mạng trên chip với các kịch bản đánh giá khác nhau. Trong mục này sẽ trình bày giải pháp đánh giá mạng trên chip tái cấu hình bằng cách kết hợp với những kịch bản đánh giá truyền thông như đã trình bày ở Mục 4.2.2 và mô phỏng các trường hợp tái cấu hình mạng trên chip.

Kịch bản cấu hình: Cùng hướng đến mục tiêu hoàn thiện giải pháp mạng trên chip tái cấu hình, kiến trúc đã được đề xuất cần được khảo sát đánh

giá khả năng đáp ứng truyền thông mạng trên chip trong nhiều trường hợp. Để khảo sát khả năng thích ứng với các thay đổi cấu hình, mạng trên chip tái cấu hình cần được mô phỏng đánh giá trong nhiều trường hợp cấu hình khác nhau để làm nổi rõ các vấn để tái cấu hình. Do đó, quá trình mô phỏng đánh giá cần có các kịch bản cấu hình để khảo sát khả năng thích ứng cấu hình của giải pháp cũng như hiệu quả truyền thông của giải pháp.

Trên cơ sở tái sử dụng nền tảng đánh giá ở Mục 4.2.2, tất cả các hoạt động mô phỏng truyền thông và cấu hình hệ thống đều được thực hiện qua lõi IP cho mô phỏng. Vậy nên, hoạt động mô phỏng đánh giá mạng trên chip tái cấu hình cũng thay đổi theo kịch bản được thực hiện thông qua lõi IP sau khi được bổ sung thêm khả năng điều khiển cấu hình của mạng trên chip.

Kết quả khảo sát hoạt động truyền thông mạng trên chip với cấu trúc liên kết dạng lưới hai chiều 2D-mesh đã cho thấy sự phân bố không đều của tải tin trong toàn mạng [36]. Vậy nên, trong quá trình mô phỏng sự thay đổi cấu hình cần khảo sát nhiều trường hợp bộ định tuyến ở những vị trí không đối xứng và vùng có mật độ tải tin khác nhau. Trong quá trình đánh giá này, những trường hợp bộ định tuyến bị cấm được mô phỏng ở những vị trí như sau:

- **Vị trí góc mạng** là trường hợp bộ định tuyến bị cấm ở ngay tại các góc kết cuối của cả phương dọc và phương ngang của mạng sử dụng cấu trúc liên kết dạng lưới hai chiều 2D-mesh. Những bộ định tuyến ở góc mạng chỉ liên kết với hai bộ định tuyến lân cận ở trên cả hai phương ngang và dọc.

- **Ví trí biên mạng** là trường hợp bộ định tuyến bị cấm ở kết cuối của một phương ngang hoặc phương dọc. Bộ định tuyến ở biên mạng chỉ liên kết với ba bộ định tuyến lân cận và một hướng ra sẽ bị chặn do không sử dụng trong cấu trúc liên kết dạng lưới hai chiều 2D-mesh.

- **Vị trí bên trong mạng** là những trường hợp bộ định tuyến ở phía bên trong của mạng sử dụng cấu trúc dạng lưới hai chiều 2D-mesh. Mỗi bộ định tuyến bên trong mạng luôn liên kết với bốn bộ định tuyến lân cận. Do đó, khi bộ định tuyến bị cấm bên trong mạng sẽ có ảnh hưởng lớn đến hoạt động truyền thông của toàn mạng.

Đánh giá hoạt động truyền thông

Quá trình kiểm tra đánh giá khả năng đáp ứng các yêu cầu truyền thông của hệ thống tái cấu hình luôn cần được quan tâm và thực hiện nhằm khẳng định những kết quả đạt được của giải pháp đề ra. Do vậy, hiệu năng truyền

Bảng 4.2: Thông số của mạng trên chip tái cấu hình trong mô phỏng đánh giá

Tham số	**Giá trị**
Các tham số cấu hình và mạng	
Cấu trúc liên kết	2D mesh, kích thước 5×5
Điều kiển luồng	Cơ chế Credit-based
Điều chuyển dữ liệu	Cơ chế Wormhole
Định tuyến	Giải thuật định tuyến XY tại nguồn và cập nhật định tuyến
Các tham số mô phỏng	
Mô hình tải tin	Phân bố lấy bù, tải đồng dạng
Kích thước gói tin	4 flits
Tỷ lệ tải tin	$1\% \div 90\%$
Số bộ định tuyến bị cấm	1 (inner, border, conner)

thông của giải pháp cập nhật thông tin định tuyến đã được trình bày trong Chương 3 được mô phỏng đánh giá với nhiều cấu hình mạng khác nhau. Dựa trên nền tảng đánh giá truyền thông mạng trên chip đã được phát triển [34] và kết hợp với các kịch bản tái cấu hình đã trình bày để mô phỏng đánh giá hoạt động truyền thông của giải pháp cập nhật định tuyến.

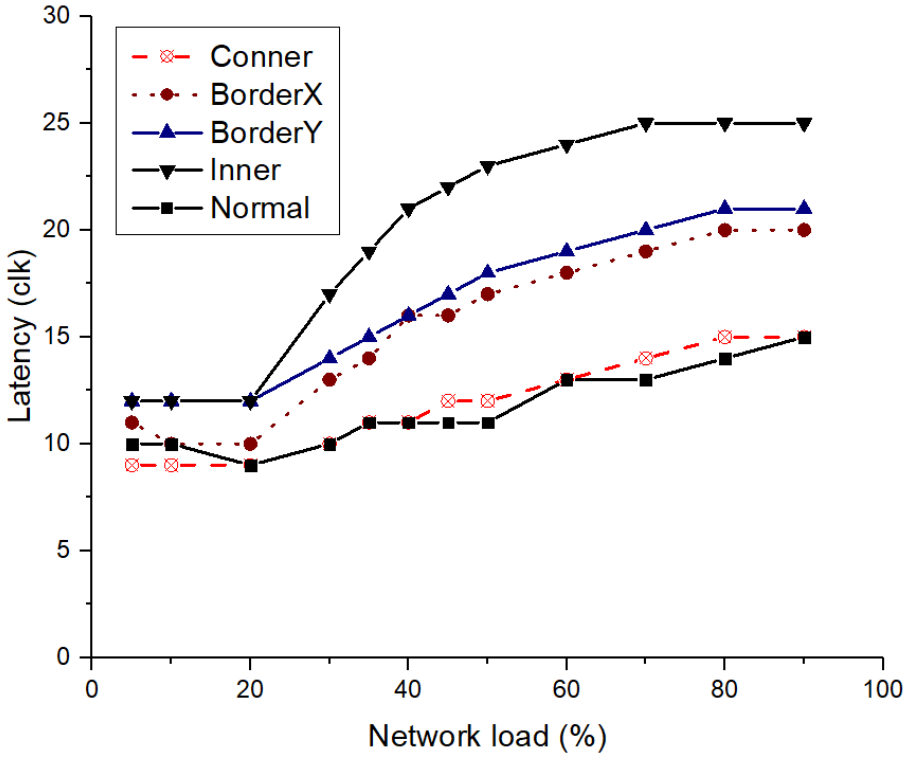
Đánh giá truyền thông của mạng trên chip tái cấu hình sử dụng các lõi IP cho đánh giá với khả năng thay đổi thông số phát tải tin theo các kịch bản đánh giá linh hoạt và ổn định [36]. Các tham số về cấu hình mạng trên chip tái cấu hình và hoạt động mô phỏng đánh giá được mô tả chi tiết trong Bảng 4.2. Kết quả đánh giá được kiểm tra và tính toán ra các thông số hiệu năng truyền thông dựa trên các tập tin dữ liệu đã được lưu trong quá trình mô phỏng hoạt động truyền thông của mạng trên chip tái cấu hình.

4.3.2 Mô phỏng và kết quả

Hoạt động mô phỏng truyền thông mạng trên chip tái cấu hình được thực hiện dựa trên nền tảng đánh giá mạng trên chip và giải pháp đánh giá đã được nên ở Mục 4.3.1. Trong đó, mạng trên chip được xây dựng từ các bộ định tuyến có khả năng cập nhật thông tin định tuyến (xem Mục 3.3.1) với cấu trúc liên kết dạng lưới hai chiều có kích thước là 5×5. Song song với quá trình mô phỏng hoạt động truyền thông, quá trình đánh giá độ tin cậy của thông tin trao đổi và tính toán các thông số hiệu năng truyền thông cũng được thực hiện một cách khách quan. Kết quả mô phỏng đánh giá truyền thông của mạng trên chip tái cấu hình khẳng định giải pháp cập nhật thông tin định tuyến đáp ứng tốt cho nhiều trường hợp thay đổi cấu hình mạng mà vẫn bảo đảm các yêu cầu trao

đổi dữ liệu tin cậy và hiệu quả. Sau đây là một số kết quả thu được khi thay đổi vị trí của bộ định tuyến bị cấm với hai thông số đánh giá là: độ trễ truyền (*latency*) và thông lượng trung bình toàn mạng (*throughput*).

Hình 4.6 là kết quả đánh giá độ trễ truyền trung bình trên toàn mạng theo tỷ lệ tải tin phát vào mạng với các vị trí của bộ định tuyến bị cấm. Để làm rõ hơn ảnh hưởng khi có bộ định tuyến rời khỏi mạng, các giá trị độ trễ được so sánh với độ trễ của mạng với cùng cấu trúc dạng lưới hai chiều và kích thước mạng nhưng không có bộ định tuyến nào bị cấm (*normal*).

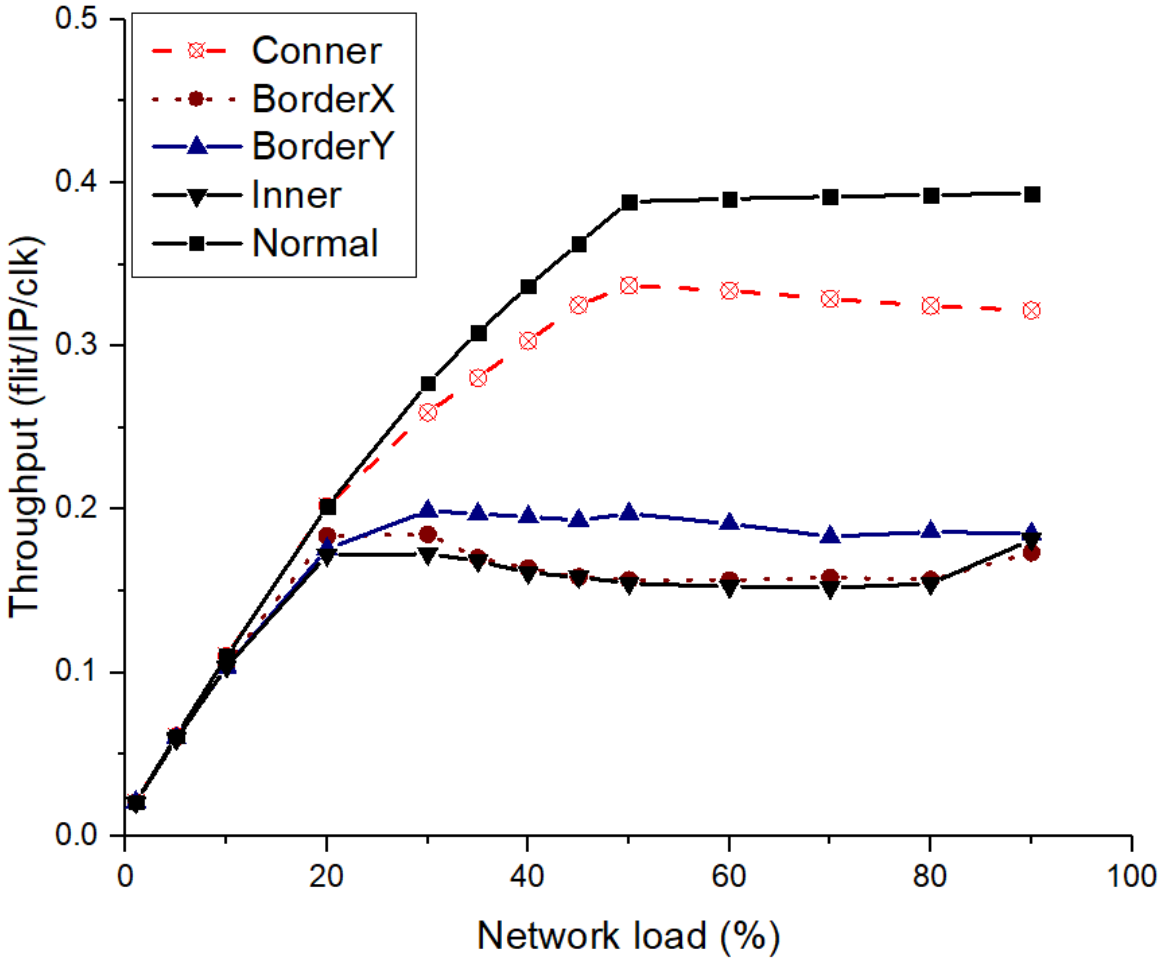


Hình 4.6: Trễ truyền tương ứng với các vị trí bị cấm.

Từ đồ thị biểu diễn độ trễ truyền (Hình 4.6) cho thấy giải pháp tái cấu hình mạng trên chip có khả năng thích ứng với nhiều sự thay đổi cấu hình mạng. Kết quả đánh giá cho thấy độ trễ truyền của mạng ổn định đối với các giá trị tải nhỏ hơn 20% và tăng nhanh khi tải tin phát vào mạng tăng lên. Kết quả ở Hình 4.6 hoàn toàn phù hợp với tính toán. Cụ thể khi không có bộ định tuyến nào bị cấm, giá trị độ trễ truyền ổn định đến giá trị $11 clk$ (tương ứng với tải mạng bằng 50%), sau đó mới tăng dần đến giá trị cực đại là $15 clk$; kết quả này tương đồng với kết quả đã có khi đánh giá mạng trên chip ở phần trước (Hình 4.4). Trong các trường hợp có bộ định tuyến bị cấm, độ trễ truyền ổn định ở giá trị $12 clk$. Sau đó,độ trễ truyền tăng nhanh và phụ thuộc vào vị trí bộ định tuyến bị chặn.

Cùng với việc đánh giá theo độ trễ truyền, giải pháp tái cấu hình còn được

đánh giá khả năng đáp ứng truyền thông với chiến lược thay đổi nhiều vị trí bộ định tuyến bị cấm. Hình 4.7 là kết quả thu được khi đánh giá khả năng truyền thông của mạng trên chip tái cấu hình với các vị trí khác nhau của bộ định tuyến bị cấm.



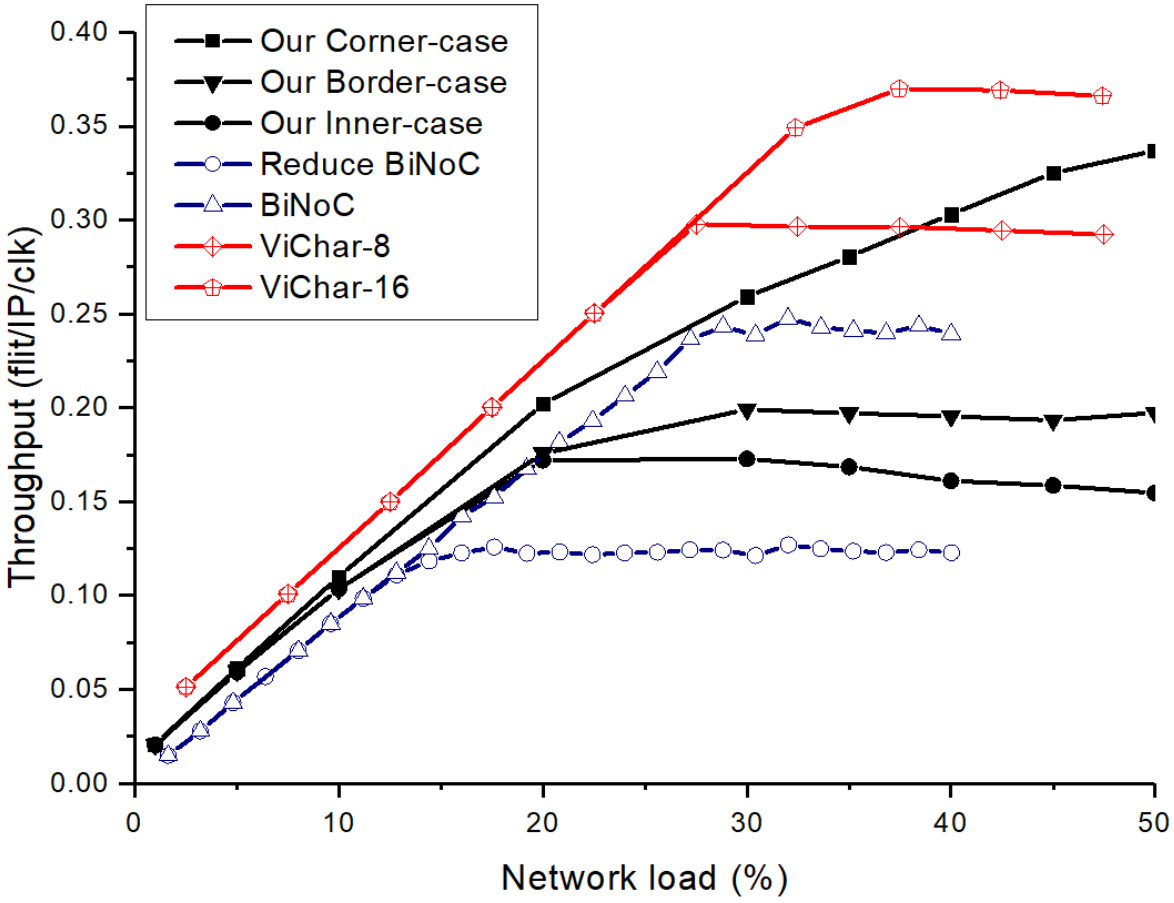
Hình 4.7: Thông lượng mạng tương ứng với các vị trí bị cấm.

Hình 4.7 cho thấy với trường hợp không bị chặn, thông lượng của mạng tăng tuyến tính theo giá trị tải tin đầu vào dưới 50%; và sau đó đạt ngưỡng bão hòa. Trong trường hợp bộ định tuyến bị cấm ở vị trí góc mạng sẽ ít ảnh hưởng đến các giao tác truyền thông của mạng nên giá trị thông lượng mạng cũng có quy luật thay đổi gần với trường hợp không bị chặn. Kết quả đánh giá cũng chỉ ra trường hợp ảnh hưởng lớn nhất của vị trí bộ định tuyến bị cấm là ở bên trong cấu trúc mạng (*inner*). Từ kết quả đánh giá thông lượng cho thấy khi vị trí bị cấm ở bên trong mạng, thông lượng bão hòa khi tải tin khoảng 20% ở giá trị $0,15\% \div 0,20\%$ (*flit/IP/clk*). Khi vị trí bộ định tuyến bị cấm ở bên trong mạng sẽ tác động đến đường định tuyến của rất nhiều giao tác truyền thông khác nhau. Điều này dẫn đến hoạt động cập nhật thông tin cho nhiều trường hợp hơn và các đường định tuyến mới cũng ảnh hưởng nhiều đến các hoạt động truyền

thông bình thường trong mạng. Qua kết quả đánh giá ở Hình 4.6 và Hình 4.7, ta nhận thấy rằng đối với trường hợp bình thường thì thông lượng chuyển sang trạng thái bão hòa khi độ trễ truyền lớn hơn $11 clk$ (tương ứng với tải mạng lớn hơn 50%), còn đối với trường hợp mạng tái cấu hình có bộ định tuyến bị cấm thì thông lượng chuyển qua trạng thái bão hòa khi độ trễ truyền lớn hơn $12 clk$ (tương ứng với tải mạng lớn hơn 20%).

Kết quả thu được trong các trường hợp trên cũng được đem so sánh với một số công trình về tái cấu hình đã được công bố như trong Hình 4.8. Trong một số công trình so sánh này cho thấy kiến trúc ViChar [50] cho kết quả thông lượng cao nhất với giá trị lớn hơn 0,36(*flit/IP/clk*) và kiến trúc Reduce BiNoC [33] có hiệu quả thông lượng thấp dưới 0,15(*flit/IP/clk*). Tuy nhiên để đạt được kết quả thông lượng cao như trên, kiến trúc ViChar phải sử dụng bộ đệm dữ liệu khá lớn ở mỗi bộ định tuyến ($80 flit$ cho ViChar-16 và $40 flit$ cho ViChar-8) và hơn nữa kiến trúc này sử dụng đến 4 kênh ảo cho mỗi kênh vật lý (Bảng 4.3). Thông lượng đạt được ở trường hợp bộ định tuyến bị chắn ở góc mạng (*conner-case*) của RNoC cho thấy thông lượng của giải pháp đề xuất gần bằng trường hợp tốt nhất của kiến trúc ViChar. Trong trường hợp xấu nhất khi bộ định tuyến bị cấm nằm bên trong mạng thì thông lượng cũng gần bằng với kiến trúc ViChar khi tải tin thấp hơn 20%. Sự so sánh toàn diện về kết quả thông lượng truyền thông và chi phí tài nguyên bên trong mỗi bộ định tuyến được trình bày trong Bảng 4.3.

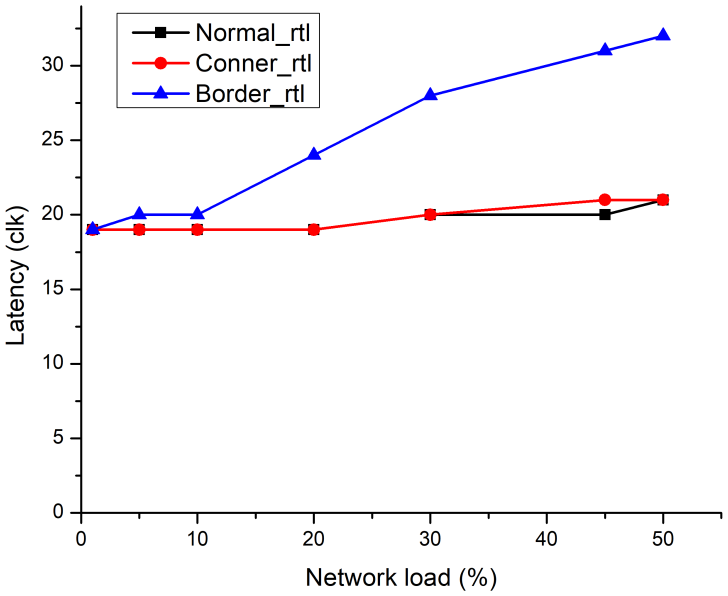
Ngay sau khi bộ định tuyến cho giải pháp tái cấu hình mạng trên chip được mô tả chi tiết ở mức dịch chuyển thanh ghi, mạng trên chip tái cấu hình được mô phỏng và đánh giá hiệu năng truyền thông trên cùng nền tảng đánh giá đã có. Kết quả đánh giá hiệu năng truyền thông được thể hiện thông qua giá trị độ trễ truyền được trình bày trong Hình 4.9. Kết quả thu được đã cho thấy được sự ảnh hưởng khác nhau giữa các trường hợp mô phỏng đáng giá khi mạng không có bộ định tuyến bị cấm (normal), bộ định tuyến bị cấm ở góc mạng (conner) và bộ định tuyến bị cấm ở biên mạng (border). Từ kết quả đánh giá thông qua độ trễ truyền thể hiện trong Hình 4.9 cho thấy quy luật thay đổi của các trường hợp bộ định tuyến bị cấm ở các vị trí khác nhau trong mạng là tương đồng với kết quả thu được khi đánh giá ở mức cao tại Hình 4.6. Tuy nhiên, kiến trúc bộ định tuyến cho giải pháp mạng trên chip tái cấu hình được xây dựng ở mức dịch chuyển thanh ghi chỉ mới hoạt động được ở chế độ đồng bộ và cần hai chu kỳ xung nhịp để truyền qua một đơn vị tin. Vậy nên, kết quả thu được khi đánh giá độ trễ truyền của mạng trên chip tái cấu hình ở Hình 4.9 có lớn hơn so với



Hình 4.8: So sánh giá trị thông lượng của các giải pháp tái cấu hình.

đánh giá ở mức cao đã thu được trước đó.

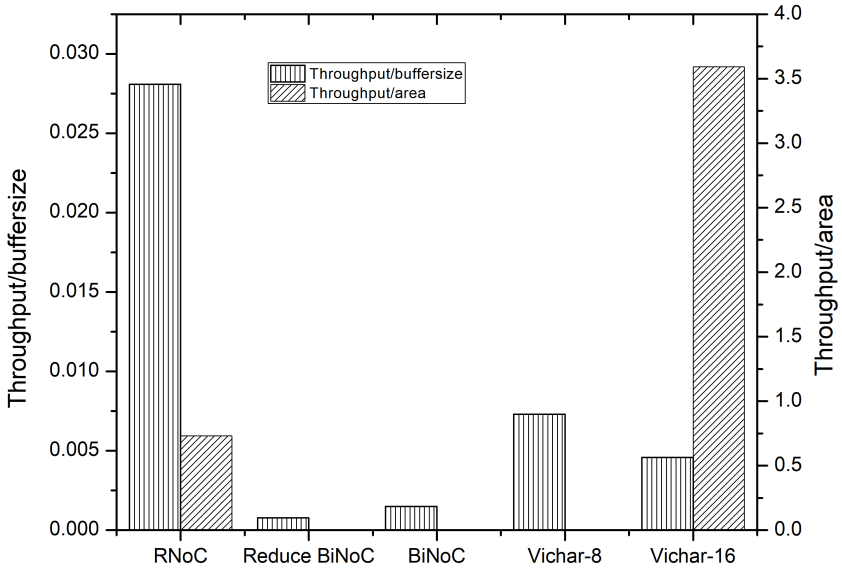
Hình 4.10 trình bày kết quả so sánh hiệu quả truyền thông của giải pháp tái cấu hình mạng trên chip trên cơ sở so sánh tỷ lệ thông lượng cực đại với tổng kích thước bộ đệm và chi phí không gian thực thi của các kiến trúc đã công bố như đã được tổng hợp trong Bảng 4.3. Kết quả này cho thấy, giải pháp được đề xuất sử dụng bộ đệm hiệu quả cao hơn nhiều so với các kiến trúc khác đã công bố. Tỷ lệ thông lượng trên mỗi đơn vị bộ đệm là 0,0129, cao gần gấp hai lần so với kiến trúc ViChar-8 và cao hơn rất nhiều so với kiến trúc BiNoC. Kết quả so sánh này cũng cho thấy bộ định tuyến trong giải pháp tái cấu hình đề xuất có công suất tiêu thụ của bộ định tuyến rất nhỏ hơn so với kiến trúc BiNoC và ViChar. Hình 4.10 cho thấy tỷ lệ thông lượng mạng lớn nhất và chi phí không gian thực thi bộ định tuyến mạng trên chip tái cấu hình còn khá lớn so với kiến trúc ViChar. Tuy nhiên, kiến trúc ViChar được thực thi bằng công nghệ TSMC $90nm$ và các đánh giá hiệu năng của ViChar chỉ cho mạng trên chip bình thường, không có trường hợp bộ định tuyến bị cấm như với đánh giá mạng trên chip tái cấu hình được đề xuất.



Hình 4.9: Độ trễ truyền của mạng trên chip tái cấu hình ở mô hình mức RTL.

Bảng 4.3: Tài nguyên truyền thông trong mỗi bộ định tuyến

Kiến trúc	**Đề xuất**	**BiNoC**	**Reduce BiNoC**	**ViChar-8**	**ViChar-16**
Tổng số bộ đệm	6	10	5	5	5
Bộ đệm/hướng	1	2	1	1	1
Kích thước bộ đệm	2*flit*	16*flit*	32*flit*	8*flit*	16*flit*
Tổng kích thước bộ đệm	12*flit*	160*flit*	160*flit*	40*flit*	80*flit*
Crossbar	$2(6 \times 6)$	10×10	5×5	5×5	5×5
Số cổng	6	5	5	5	5
Số kênh ảo	2	2	1	4	4
Công nghệ chế tạo	CMOS 130nm	TSMC 90nm	TSMC 90nm	TSMC 90nm	TSMC 90nm
Chi phí không gian	$461345\mu m^2$	48666gates	NA	NA	$101899\mu m^2$



Hình 4.10: So sánh tỷ lệ thông lượng với kích thước bộ đệm và chi phí không gian thực thi của các kiến trúc tái cấu hình.

4.4 Kết luận chương

Chương này đã tập trung trình bày thiết kế một nền tảng cho đánh giá hoạt động truyền thông mạng trên chip và áp dụng cho cả thiết kế mạng trên chip tái cấu hình với khả năng thay đổi thông số đánh giá linh hoạt, được công bố tại công trình [C2]. Nền tảng đánh giá được đề xuất gồm hai phần chính: mô phỏng hoạt động truyền thông của các thiết kế mạng trên chip dựa trên thông số cấu hình cho phép thực hiện các kịnh bản đánh giá đã xác định; chức năng đánh giá hiệu quả truyền thông dựa trên các dữ liệu thu được từ quá trình mô phỏng hoạt động truyền thông một cách khách quan. Nền tảng đánh giá được đề xuất sử dụng hai thông số đánh giá chính là độ trễ truyền và thông lượng mạng để làm rõ khả năng truyền thông của các thiết kế mạng trên chip trong các bối cảnh mô phỏng đánh giá khác nhau. Việc sử dụng dữ liệu mô phỏng để thực hiện đánh giá cho phép quá trình đánh giá hoạt động độc lập với quá trình mô phỏng, nhưng vẫn bảo đảm được độ tin cậy cao để cho ra những kết quả đánh giá khách quan. Kết quả đánh giá đã được công bố trong công trình [C3] của Hội nghị quốc tế về các Công nghệ tiên tiến trong truyền thông (ATC) diễn ra tại Hà Nội năm 2012.

Với mục tiêu đánh giá giải pháp tái cấu hình mạng trên chip đã được đề cập ở Chương 3, nền tảng đánh giá được bổ sung chức năng thay đổi cấu hình

trên chip trong quá trình mô phỏng hoạt động truyền thông nhưng vẫn bảo đảm tính khách quan của hoạt động đánh giá. Do đó, hoạt động đánh giá mạng trên chip tái cấu hình đã được thực hiện một cách linh hoạt và tin cậy với nhiều kịch bản tái cấu hình và mô phỏng hoạt động truyền thông phù hợp. Kết quả đánh giá mạng trên chip tái cấu hình được thể hiện đa dạng thông qua sự thay đổi vị trí của bộ định tuyến bị cấm, đồng thời quá trình đánh giá đã so sánh hiệu quả hoạt động truyền thông của những trường hợp tái cấu hình và không tái cấu hình để làm rõ hơn ảnh hưởng của các trường hợp tái cấu hình đến hoạt động truyền thông của hệ thống.

Kết quả đánh giá cho thấy hiệu quả truyền thông của giải pháp tái cấu hình mạng trên chip có sự khác nhau tương ứng với các vị trí của bộ định tuyến bị cấm, nhưng vẫn bảo đảm cho hoạt động truyền thông ngay cả trong những trường hợp tải tin liên tục yêu cầu phát tin (tải tin 100%). Kết quả đánh giá thiết kế mạng trên chip tái cấu hình đã được chúng tôi công bố trong công trình [J1] trên Tạp chí Điện tử-Truyền thông JEC của Hội Vô tuyến-Điện tử Việt Nam.

Kết luận và hướng phát triển

Kết luận

Việc sử dụng mô hình truyền thông mạng trên chip là một định hướng nghiên cứu đã và đang được các nhà khoa học quan tâm. Tuy nhiên, để mô hình mạng trên chip sớm được ứng dụng một cách đại trà, các nhà khoa học cần tiếp tục hoàn thiện mô hình này, đặc biệt là tăng khả năng thích ứng của mô hình đối với những thay đổi của vi mạch trong quá trình sản xuất cũng như trong quá trình hoạt động. Do vậy, chúng tôi đã nghiên cứu và đề xuất một giải pháp tái cấu hình mạng truyền thông trên chip thông qua việc đưa ra cơ chế cập nhật thông tin định tuyến. Giải pháp cập nhật thông tin định tuyến này cho phép mạng trên chip thay đổi đường định tuyến một cách linh hoạt, ứng với các thay đổi của mạng khi có bộ định tuyến bị lỗi (hay rời khỏi mạng).

Luận án này đã tổng quan hoạt động truyền thông trên chip sử dụng mô hình mạng, đặc biệt là các kiến trúc truyền thông có khả năng tái cấu hình. Để khảo sát hoạt động truyền thông mạng trên chip, tôi đã xây dựng mạng trên chip với các đặc điểm truyền thông của mạng chuyển mạch gói và định tuyến XY tại nguồn. Kết quả khảo sát mô hình mạng trên chip đã được công bố tại công trình [C1]. Hướng trọng tâm nghiên cứu giải pháp tái cấu hình trong nghiên cứu thiết kế hệ thống, luận án cũng đã hệ thống lại các nghiên cứu tái cấu hình và các giải pháp tái cấu hình. Tác giả và nhóm nghiên cứu của mình đã công bố công trình tổng quan về xu thế tái cấu hình công bố tại công trình [J2] đã được chấp nhận đăng trên Tạp chí Điện tử - Viễn thông JEC năm 2017.

Trên cơ sở nghiên cứu hoạt động truyền thông và tái cấu hình mạng trên chip, luận án đã đề xuất giải pháp cập nhật thông tin định tuyến cho mạng trên chip tái cấp hình khi có bộ định tuyến rời khỏi mạng:

- Giải pháp cập nhật định tuyến cho phép thay đổi đường đi của thông tin linh hoạt để thích ứng với các thay đổi cấu hình mạng trên chip ngay cả

khi hệ thống đang hoạt động. Đồng thời giải pháp này vẫn giữ được ưu điểm của hoạt động định tuyến tĩnh tại nguồn cho các giao tác truyền thông không tái cấu hình để bảo đảm được hiệu quả truyền thông tối ưu cho toàn hệ thống. Tiếp đó, luận án đã phát triển kiến trúc bộ định tuyến có khả năng tái cấu hình nhằm thực thi giải pháp nêu trên. Tiếp đó, kiến trúc này được mô hình hóa để đánh giá hoạt động truyền thông tái cấu hình. Kiến trúc bộ định tuyến đã được mô hình hóa bằng ngôn ngữ phần cứng VHDL ở mức chuyển dịch thanh ghi (RTL: *Register Transfer Level*) và thực thi với công nghệ CMOS $130nm$ của hãng Global Foundry. Kết quả thực nghiệm cho thấy rõ tính khả thi của kiến trúc đã được xây dựng hoạt động ở tần số $100MHz$ với chi phí không gian thực thi $461345\mu m^2$ và công suất tiêu thụ thấp ($7{,}58mW$). Các kết quả này đã được công bố trên tạp chí khoa học chuyên ngành JEC ([J1]).

- Cuối cùng, trong luận án này, tôi cũng phát triển phương pháp và nền tảng mô hình hóa, mô phỏng đa lớp sử dụng các ngôn ngữ mô tả phần cứng khác nhau (C++, SystemC, VHDL) để xây dựng nền tảng đánh giá hoạt động truyền thông mạng trên chip. Nền tảng đề xuất cho phép mô phỏng và đánh giá nhanh hiệu năng truyền thông mạng trên chip với các kịch bản đánh giá đa dạng cho cả mạng trên chip thông thường và mạng trên chip tái cấu hình. Các kết quả này đã được công bố trên kỷ yếu hội nghị khoa học quốc tế (ATC, ICDV) và tạp chí khoa học chuyên ngành JEC ([C2], [C3], và [J1]).

Hướng phát triển

Sau thời gian nghiên cứu các vấn đề liên quan đến hoạt động truyền thông mạng trên chip và thực hiện giải thuật cập nhật định tuyến sử dụng cho mạng trên chip tái cấu hình đã thu được nhiều kết quả khả quan như đã trình bày ở trên. Giải pháp này cũng cần được nghiên cứu phát triển một cách toàn diện hơn nhằm tiến đến quá trình tổng hợp kiến trúc lên vi mạch cụ thể; đồng thời giải pháp cũng có thể phát triển để tối ưu hơn thông qua việc xây dựng các cơ chế điều khiển truyền thông ở các mức cao hơn.

Với thời gian thực hiện đề tài không dài cho một định hướng nghiên cứu thiết kế vi mạch, giải pháp cập nhật định tuyến đã được đề xuất và xây dựng mô hình hoạt động và mô tả ở mức khả thi thông qua kết quả tổng hợp đánh giá đã được trình ở Mục 3.4, nhưng vẫn chưa được thực hiện trên vi mạch cụ

thể. Hơn nữa, việc tổng hợp trên vi mạch cụ thể cần được thực hiện kết hợp với các lõi IP để hoàn thiện chức năng cụ thể; lúc đó chúng ta mới có đầy đủ cơ sở đánh giá hiệu quả của hoạt động tái cấu hình một cách chi tiết hơn.

Bám sát mục tiêu đề ra ngay từ đầu của đề tài là nghiên cứu hoạt động truyền thông sử dụng mạng trên chip tái cấu hình, đề tài đã đưa ra giải pháp tái cấu hình bằng cách thay đổi thông tin định tuyến ngay tại bộ định tuyến đã cho kết quả hoạt động rất khả quan. Tuy nhiên, giải pháp xử lý phân tán tại các bộ định tuyến chỉ có khả năng thích ứng các hoạt động truyền thông cho những vị trí lân cận của bộ định tuyến bị cấm. Điều này có thể làm thay đổi tải tin cục bộ và ảnh hưởng đến khả năng cân bằng tải tin của toàn mạng. Do vậy, một giải pháp kết hợp khả năng xử lý linh hoạt tại bộ định tuyến để thích ứng tức thời với sự thay đổi cấu hình mạng kết hợp với cơ chế điều khiển truyền thông ở mức cao hơn sẽ cho phép hệ thống hoạt động ổn định hơn và thu được nhiều kết quả khả quan. Với định hướng phát triển này chúng tôi bước đầu đã đề xuất được một cơ chế theo dõi và cập nhật đường định tuyến. Kết quả ban đầu được công bố tại hội nghị IEEE ICICDT 2016 (công trình [C5]).

Danh mục công trình khoa học của tác giả liên quan đến luận án

C1 Nam-Khanh Dang, **Thanh-Vu Le-Van**, Xuan-Tu Tran (2011), "FPGA Implementation of a Low Latency and High Throughput Network-on-Chip Router Architecture", *in Proceeding of the 2011 International Conference on Integrated Circuits, Design, and Verification (ICDV 2011)*, pp. 112-116, Hanoi, August 2011, ISBN: 978-4-88552-258-1.

C2 **Thanh-Vu Le-Van**, Dien-Tap Ngo, Xuan-Tu Tran (2012), "A SystemC based Simulation Platform for Network-on-Chip Architectures", *in Proceeding of the 2011 International Conference on Integrated Circuits, Design, and Verification (ICDV 2012)*, pp. 132-136, Danang, August 2012, ISBN: 978-4-88552-264-2.

C3 **Thanh-Vu Le Van**, Xuan-Tu Tran (2012), "Simulation and Performance Evaluation of a Network-on-Chip Architecture based on SystemC", *in Proceeding of the 2012 International Conference on Advanced Technologies for Communications (ATC 2012)*, pp. 170-175, Hanoi, Vietnam, October 2012, ISBN: 978-1-4673-4350-3.

C4 Thi-Thuy Nguyen, **Thanh-Vu Le-Van**, Hung K. Nguyen, Xuan-Tu Tran (2016), "Routing-path Tracking and Updating Mechanism in Reconfigurable Network-on-Chips", *in Proceeding of the 2016 IEEE International Conference on Integrated Circuits Design and Technology (ICICDT 2016)*, 27-29 June 2016, Ho Chi Minh city, Vietnam.

J1 **Thanh-Vu Le-Van**, Xuan-Tu Tran (2014), "High-Level Modeling and Simulation of a Novel Reconfigurable Network-on-Chip Router", *REV Journal on Electronics and Communications (JEC)*, pp. 68-74, Vol. 4, No. 3-4, July-December, 2014, ISSN: 1859-387X.

J2 Hung K. Nguyen, **Thanh-Vu Le-Van**, Xuan-Tu Tran (2017). "A Survey on Reconfigurable System-on-Chips". *REV Journal on Electronics and Communications (JEC)*, ISSN: 1859–387X (Accepted)

Tài liệu tham khảo

[1] Ankur Agarwal, Cyril Iskander, and Ravi Shankar. Survey of Network on Chip (NoC) Architectures & Contributions. *Engineering, Computing and Architecture*, 3(1), 2009.

[2] E Beigne, Fabien Clermidy, Pascal Vivet, A Clouard, and M Renaudin. An Asynchronous NoC Architecture Providing Low Latency Service and its Multi-level Design Framework. In *proceedings the 11th IEEE International Symposium on Asynchronous Circuits and Systems (ASYNC'05)*, pages 54 – 63, 2005.

[3] Luca Benini and Giovanni De Micheli. Networks on Chips: A New SoC Paradigm. *IEEE Computer*, 35(1):70–78, January 2002.

[4] Davide Bertozzi and Luca Benini. Xpipes: a network-on-chip architecture for gigascale systems-on-chip. In *Circuits and Systems Magazine, IEEE*, volume 4, pages 18–31, 2004.

[5] Tobias Bjerregaard and Shankar Mahadevan. A Survey of Research and Practices of Network-on-Chip. In *ACM Computing Surveys*, volume 38, 2006.

[6] Christophe Bobda, Mateusz Majer, Dirk Koch, Ali Ahmadinia, and Jurgen Teich. A Dynamic NoC Approach for Communication in Reconfigurable Devices. In *International Conference on Field-Programmable Logic and Applications (FPL), ser. Lecture Notes in Computer Science (LNCS)*, volume 3203, pages 1032–1036, 2004.

[7] Chritophe Bobda, Ali Ahmadinia, Mateusz Majer, Jurgen Teich, Sandor Fekete, and Jan van der Veen. DyNoc: A dynamic infrastructure for communication in dynamically reconfugurable devices. In *proceedings of Field Programmable Logic and Applications*, 2005.

[8] Evgeny Bolotin, Israel Cidon, and Ran Ginosarand Avinoam Kolodny. QNoC: QoS Architecture and Design Process for Network on Chip. *Journal of Systems Architecture: the EUROMICRO Journal - Special issue: Networks on chip*, 50:105–128, 2004.

[9] Sao-Jie Chen, Ying-Cherng Lan, Wen-Chung Tsai, and Yu-Hen Hu. *Reconfigurable Networks-on-Chip*. Springer, 2012.

[10] Ge-Ming Chiu. The Odd-Even Turn Model for Adaptive Routing. *IEEE Transactions on Parallel and Distributed Systems*, 11(7):729–738, 2000.

[11] Caroline Concatto, Debora Matos, Luigi Carro, Fernanda Kastensmidt, Altamiro Susin, and Marcio Kreutz. NoC Power Optimization Using a Reconfigurable Router. In *proceedings of the 2009 IEEE Computer Society Annual Symposium on VLSI (ISVLSI '09)*, 2009.

[12] Rachid Dafali, Jean-Philippe Diguet, and Marc Sevaux. Key Research Issues for Reconfgurable Network-on-Chip. In *proceedings of the International Conference on Reconfigurable Computing and FPGAs*, 2008.

[13] William J. Dally and Brian Towles. Route packets, not wires: on-chip interconnection networks. In *proceedings of the 38th annual Design Automation Conference (DAC'01)*, pages 684 – 689, New York, NY, USA, 2001.

[14] William James Dally and Brian Towles. *Principles and Practices of Interconnection Networks*. Morgan Kaufmann Publishers, 2004.

[15] Nam-Khanh Dang, Thanh-Vu Le-Van, and Xuan-Tu Tran. Fpga implementation of a low latency and high throughput network-on-chip router architecture. In *The Institute of Electronics, Information and Communication Engineers*, 2011.

[16] Carl Ebeling, Darren C. Cronquist, and Paul Franklin. RaPiD - Reconfigurable Pipelined Datapath. In *proceedings of the 6th International Workshop on Field-Programmable Logic and Applications*, 1996.

[17] Mohammad Abdullah Al Faruque, Thomas Ebi, and Jorg Henkel. ROAdNoC: Runtime Observability for an Adaptive Network on Chip Architecture. In *proceedings of the Computer-Aided Design - ICCAD 2008*, 2008.

[18] Jose Flich and Jose Duato. Logic-Based Distributed Routing for NoCs. *IEEE Computer Architecture*, 7(1):13–16, 2008.

[19] Jose Flich, A. Mejia, P. Lopez, and Jose Duato. Region-Based Routing: An Efficient Routing Mechanism to Tackle UnreliableHardware in Network on Chips. In *proceedings of the first International Symposium on Networks-on-Chip (NOCS)*, 2007.

[20] Jose Flich, S. Rodrigo, and Jose Duato. An Efficient Implementation of Distributed Routing Algorithms for NoCs. In *proceedings of the Second ACM/IEEE International Symposium on Networks-on-Chip*, 2008.

[21] Philip Garcia, Katherine Compton, Michael Schulte, Emily Blem, and Wenyin Fu. An Overview of Reconfigurable Hardware in Embedded Systems. *EURASIP Journal on Embedded Systems*, 2006(1):13–13, January 2006.

[22] Samuel Garcia and Bertrand Granado. Ollaf: A fine grained dynamically reconfigurable architecture for os support. *EURASIP Journal on Embedded Systems*, 2009(1):574716, Nov 2009.

[23] Christopher J. Glass and Lionel M. Ni. The Turn Model for Adaptive Routing. In *proceedings of the International Symposium on Computer Architecture*, pages 278–287, 1992.

[24] Cristian Grecu, Partha Pande, Andre Ivanov, and Res Saleh. BIST for Network-on-Chip Interconnect Infrastructures. In *VLSI Test Symposium, 2006. Proceedings. 24th IEEE*, 2006.

[25] Pierre Guerrier and Alain Greiner. A Generic Architecture for On-Chip Packet-Switched Interconnections. In *proceedings of the Conference Design, Automation and Test in Europe Conference and Exhibition (DATE '00)*, pages 250–256, Paris, 2000.

[26] Gu Haiyun. Survey of Dynamically Reconfgurable Network-on-Chip. In *proceedings of Future Computer Sciences and Application (ICFCSA)*, 2011.

[27] Jingcao Hu and Radu Marculescu. DyAD - Smart Routing for Networks-on-Chip. In *proceedings of the ACM/IEEE Design Automation Conference*, pages 260–263, 2004.

[28] Nguyen K. Hung, Peng Cao, Xue-Xiang, Jun Yang, Longxing Shi, Min Zhu, Liebo Liu, and Shaojun Wei. Hardware Software Co-design of H.264 Baseline Encoder on Coarse-Grained Dynamically Reconfigurable Computing System-on-Chip. *The Institute of Electronics, Information and Communiction Engineers*, 2013.

[29] Natalie Enright Jerger and Li-Shiuan Peh. On-Chip Networks. In *Synthesis Lectures on Computer Architecture*, 8. Morgan & Claypool, 2009.

[30] Lionel Ni Jose Duato, Sudhakar Yalamanchili. *Interconnection Networks: An Engineering Approach*. Morgan Kaufmann Publishers, 2003.

[31] Slavisa Jovanovic, Camel Tanougast, Chritophe Bobda, and Serge Weber. CuNoC: A dynamic scalable communication structure for dynamically reconfigurable FPGAs. *Microprocessors and Microsystems*, 33(1):24–36, 2009.

[32] Oscar J. Kuiken, Xiao Zhang, and Hans G. Kerkhoff. Built-In Self-Diagnostics for a NoC-Based Reconfigurable IC for Dependable Beamforming Applications. In *proceedings of the IEEE International Symposium on Defect and Fault Tolerance of VLSI Systems*, pages 45 – 53, Boston, MA, 2008.

[33] Ying-Cherng Lan, Shih-Hsin Lo, Yueh-Chi Lin, Yu-Hen Hu, and Sao-Jie Chen. BiNoC: A Bidirectional NoC Architecture with Dynamic Self-Reconfigurable Channel. In *proceedings of the 2009 3rd ACM/IEEE International Symposium on Networks-on-Chip*, 2009.

[34] Thanh-Vu Le-Van, Dien-Tap Ngo, and Xuan-Tu Tran. A SystemC based Simulation Platform for Network-on-Chip Architectures. In *proceedings of the Institute of Electronics, Information and Communication Engineers*, 2012.

[35] Thanh-Vu Le-Van and Xuan-Tu Tran. High-Level Modeling and Simulation of a Novel Reconfigurable Network-on-Chip Router. *REV Journal on Electronics and Communications*, 4, 2014.

[36] Thanh-Vu Le-Van, Xuan-Tu Tran, and Dien-Tap Ngo. Simulation and Performance Evaluation of a Network-on-Chip Architecture based on SystemC. In *proceedings of the International Conference on Advanced Technologies for Communications (ATC'12)*, pages 170–175, 2012.

[37] Jean-Jacques Lecler and Gilles Baillieu. Application driven network-on-chip architecture exploration & refinement for a complex soc. *Design Automation for Embedded Systems*, 15(2):133–158, June 2011.

[38] Dee Lin and Kurt Shuler. Optimizing Enterprise-Class SSD Host Controller Design with Arteris FlexNoC Network-On-Chip Interconnect IP. Technical report, Arteris, Inc, 2016.

[39] ARM Ltd. *ARM CoreLink CCI-550 Cache Coherent Interconnect*, 2015.

[40] Zhonghai Lu, Rikard Thid, Mikael Millberg, Erland Nilsson, and Axel Jantsch. NNSE: Nostrum Network-on-Chip Simulation Environment. In *proceedings of the Swedish System-on-Chip Conference (SSoCC'03)*, 2005.

[41] Terrence Mak, Peter Y. K. Cheung, Kai-Pui Lam, and Wayne Luk. Adaptive Routing in Network-on-Chips Using a Dynamic-Programming Network. *IEEE Transactions on Industrial Electronics*, 58(8):3701–3716, 2011.

[42] Ran Manevich, Isreal Cidon, Avinoam Kolodny, and Isask'har Walter. Centralized Adaptive Routing for NoCs. *Computer Architecture Letters*, 9(2):57 – 60, 2010.

[43] Radu Marculescu, Umit Y. Ogras, Li-Shiuan Peh, Natalie Enright Jerger, and Yatin Vasant Hoskote. Outstanding Research Problems in NoC Design: System, Microarchitecture, and Circuit Perspectives. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 28(1):3–21, 2009.

[44] B. Mei, M. Berekovic, and J-Y. Mignolet. *ADRES & DRESC: Architecture and Compiler for Coarse-Grain Reconfigurable Processors*, pages 255–297. Springer Netherlands, 2007.

[45] A. Mejia, Jose Duato, Jose Flich, Sven-Arne Reinemo, and Tor Skeie. Segment-Based Routing: An Efficient Fault-Tolerant Routing Algorithm for Meshes and Tori. In *proceedings of the Parallel and Distributed Processing Symposium (IPDPS 2006)*, 2006.

[46] Andres Mejia, Jose Flich, and Jose Duato. On the Potentials of Segment-Based Routing for NoCs. In *proceedings of the 37th International Conference on Parallel Processing*, pages 594–603, 2008.

[47] Kullai Reddy Meka and Jatindra Kumar Deka. Demand Based Routing in Network-on-Chip(NoC). In *proceeding in the 2011 International Conference on Embedded Systems and Applications (ESA'11)*, 2011.

[48] George Michelogiannakis, Nan Jiang, Daniel Becker, and William J. Dally. Packet Chaining: Efficient Single-Cycle Allocation for On-Chip Networks. *IEEE Computer Architecture Letters*, 10(2):33–36, 2011.

[49] Yasuyuki Miura, Kentaro Shimozono, Naohisa Fukase, and Shigeyoshi Watanabe. An Adaptive Routing Algorithm of 2-D Torus Network Based on Turn Model: The Communication Performance. *International Journal of Networking and Computing*, 5(1):223–238, 2015.

[50] Chrysostomos A. Nicopoulos, Dongkook Park, Jongman Kim, and Chita R. Das N. Vijaykrishnan, Mazin S. Yousif. ViChaR: A Dynamic Virtual Channel Regulator for Network-on-Chip Routers. In *proceedings of the 39th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO'06)*, 2006.

[51] Vincent Nollet, Theodore Marescaux, Prabhat Avasare, D. Verkest, and Jean-Yves Mignolet. Centralized Run-Time Resource Management in a Network-on-Chip Containing Reconfigurable Hardware Tiles. In *proceedings of the Design, Automation and Test in Europe (DATE '05)*, volume 1, pages 234–239, 2005.

[52] Vincent Nollet, Thesodore Marescaux, and Diederik Veerkest. Operating-System Controlled Network on Chip. In *rroceedings of the 41st annual Design Automation Conference (DAC '04)*, 2004.

[53] Partha Pratim Pande and Andre Ivanov. Performance Evaluation and Design Trade-Offs for Network-on-Chip Interconnect Architectures. *IEEE Transactions on Computers*, 54(8):1025–1040, 2005.

[54] Thilo Pionteck, Carsten Albrecht, and Roman Koch. A Dynamically Reconfigurable Packet-Switched Network-on-Chip. In *proceedings of the Design, Automation and Test in Europe (DATE'06)*, volume 1, 2006.

[55] Thilo Pionteck, Roman Koch, and Carsten Albrecht. Applying Partial Reconfiguration to Networks-On-Chips. In *proceedings of the International Conference on Field Programmable Logic and Applications (FPL'06)*, pages 1–6, 2006.

[56] Amir-Mohammad Rahmani, Ali Afzali-Kusha, and Massoud Pedram. NED: A Novel Synthetic Traffic Pattern for Power/Performance Analysis of Network-on-chips Using Negative Exponential Distribution. *Journal of Low Power Electronics, American Scientific Publishers*, 5(3):396–405, 2009.

[57] ITU-T X-Series Recommendations. Information Technology - Open Systems Interconnection - Basic reference model: The Basic model. In *www.itu.int/rec*, 1994.

[58] Samuel Rodrigo, Jose Flich, Jose Duato, and Mark Hummel. Efficient Unicast and Multicast Support for CMPs. In *proceedings of the 41st annual IEEE/ACM International Symposium on Microarchitecture (MICRO 41)*, pages 364–375, 2008.

[59] Samuel Rodrigo, Jose Flich, A. Roca, S. Medardoni, D. Bertozzi, J. Camacho, F. Silla, and Jose Duato. Addressing Manufacturing Challenges with Cost-Efficient Fault Tolerant Routing. In *proceeding of the 2010 Fourth ACM/IEEE International Symposium on Networks-on-Chip (NOCS)*, pages 25–32, 2010.

[60] Hartej Singh, Ming-Hau Lee, Guangming Lu, Fadi J. Kurdahi, and Nader Bagherzadeh. MorphoSys: an integrated reconfigurable system for data-parallel and computation-intensive applications. *IEEE Transactions on Computers*, 49(5):465–481, 2000.

[61] Mikkel Bystrup Stensgaard and Jens Spars. ReNoC: A Network-on-Chip Architecture with Reconfigurable Topology. In *proceedings of the Second ACM/IEEE International Symposium on Networks-on-Chip*, 2008.

[62] Matthias Bo Stuart, Mikkel Bystrup Stensgaard, and Jens Spars. The ReNoC Reconfigurable Network-on-Chip: Architecture, Configuration Algorithms, and Evaluation. In *ACM Transactions on Embedded Computing Systems (TECS'11)*, volume 10, 2011.

[63] Suboh Suboh, Mohamed Bakhouya, Jaafar Gaber, and Tarek El-Ghazawi. Analytical Modeling and Evaluation of Network-on-Chip Architectures. In *proceedings of the High Performance Computing and Simulation (HPCS)*, pages 615–622, 2010.

[64] XPP Technologies. Reconfiguration on XPP-III Processors. In *White Paper*, 2006.

[65] Vaishali Tehre and Ravindra Kshirsagar. Survey on Coarse Grained Reconfigurable Architectures. *International Journal of Computer Applications*, 48(16), 2012.

[66] G. Theodoridis, Dimitrios Soudris, and Stamatis Vassiliadis. *Fine- and Coarse-grain Reconfigurable Computing*, chapter 2, pages 89–149. Springer Netherlands, 2007.

[67] Xuan-Tu Tran, editor. *Emerging Aspects in Electronics and Communication Engineering*, chapter 4, pages 10–20. Vietnam National University, 2013.

[68] Xuan-Tu Tran, Yvain Thonnart, Jean Durupt, Vincent Beroulle, and Chantal Robach. Design-for-test approach of an asynchronous network-on-chip architecture and its associated test pattern generation and application. *IET Journal on Computers and Digital Techniques*, 3:487–500, 2009.

[69] Nikolaos S. Voros, Michael Hübner, Jürgen Becker, Matthias Kühnle, Florian Thomaitiv, Arnaud Grasset, Paul Brelet, Philippe Bonnot, Fabio Campi, Eberhard Schüler, Henning Sahlbach, Sean Whitty, Rolf Ernst, Enrico Billich, Claudia Tischendorf, Ulrich Heinkel, Frank Ieromnimon, Dimitrios Kritharidis, Axel Schneider, Joachim Knaeblein, and Wolfram Putzke-Röming. MORPHEUS: A heterogeneous dynamically reconfigurable platform for designing highly complex embedded systems. *ACM Transactions on Embedded Computing Systems*, 12(3), 2013.

[70] Cesar Albenes Zeferino and Altamiro Amadeu Susin. SoCIN: A Parametric and Scalable Network-on-Chip. In *proceedings of the 16th Symposium on Integrated Circuits and Systems Design (SBCCI'03)*, 2003.

[71] Lei Zhang, Yinhe Han, Qiang Xu, Xiao wei Li, and Huawai Li. On Topology Reconfiguration for Defect-Tolerant NoC-Based Homogeneous Manycore Systems. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 17(9):1173 – 1186, 2009.

[72] Lei Zhang, Yue Yu, Jianbo Dong, Yinhe Han, Shangping Ren, and Xiaowei Li. Performance asymmetry aware topology Virtualization for Defect-tolerant NoC-based Many-core Processors. In *proceeding of the Design, Automation & Test in Europe Conference & Exhibition (DATE), 2010*, pages 1566–1571, 2010.

[73] Zhen Zhang, Alain Greiner, and Sami Taktak. A Reconfigurable Routing Algorithm for a Fault-Tolerant 2D-Mesh Network-on-Chip. In *proceedings of the 45th annual Design Automation Conference (DAC'08)*, pages 441–446, 2008.

[74] Jia Zhao, Sailaja Madduri, Ramakrishna Vadlamani, Wayne Burleson, and Russell Tessier. A Dedicated Monitoring Infrastructure for Multicore Processors. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 19(6):1011–1022, 2011.

[75] Jingyang Zhu, Zhiliang Qian, and Chi ying Tsui. BiLink: A High Performance NoC Router Architecture using Bi-directional Link with Double Data Rate. *the VLSI Journal Integration*, (55):30–42, 2016.