WeDBless : Weighted Deflection Bufferless Router for Mesh NoCs

Simi Zerine Sleeba
Model Engineering College
Kochi, India
simi@mec.ac.in

John Jose
Rajagiri School of Engineering
& Technology, Kochi, India
johnj@rajagiritech.ac.in

Mini M.G.
Model Engineering College
Kochi, India
mininair@mec.ac.in

ABSTRACT

Bufferless NoC routers employing deflection routing are gaining popularity due to their power and area efficiency. We propose WeDBless, a bufferless deflection router that reduces deflection rate of flits by employing port allocation based on weighted deflection of flits. The proposed method directs the frequently misrouted flits towards their destination by increasing their probability of getting a productive output port. Our evaluations on synthetic traffic patterns show that WeDBless achieves significant reduction in deflection rate, average flit latency and improvement in network saturation point compared to the state-of-the-art bufferless router with reduced complexity in route computing logic.

Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design**]: Network Communication

Keywords

Deflection rate; output port selection; latency reduction

1. INTRODUCTION & BACKGROUND

Efficient microarchitecture and cost effective routing algorithms are highly essential characteristics for NoC routers. Traditional virtual channel routers (VCR) employ buffers in the input ports. Buffers contribute significantly to dynamic and static power [1, 3]. With an aim of reducing chip area and power, bufferless deflection routers are introduced [3]. In a bufferless deflection router, all flits arriving at the input ports have to pass through one of the output ports at the end of the pipeline cycle. This can lead to misrouting of flits and hence can increase the latency of flits. Effective output port selection in deflection routers is a critical design issue.

The baseline bufferless router, BLESS [3] uses an age based flit ranking scheme for output port selection. Because of the sequential port allocation scheme, the router pipeline

latency in BLESS is high leading to lower operating frequency of the network. This performance issue in BLESS is addressed in CHIPPER [2] by parallel port allocation. CHIPPER makes sure that the highest priority flit(golden flit) is assigned the productive port in every router. But it leads to increase in deflection rate compared to BLESS since all non-golden flits are assigned random ports. In this paper, we briefly describe the working of Weighted Deflection Bufferless (WeDBless) router and analyse the experimental results.

2. WEDBLESS ROUTER

We introduce a novel routing algorithm for bufferless deflection routers which prioritises flits based on Weighted Deflection Count(WDC) and assigns output ports based on Directional Weights (DW). WDC and DWs are computed in each router and are incorporated in the flit itself. A small side buffer which buffers one ejection ready flit is also provided to reduce deflections. Reduction in the deflection rate leads to reduction of dynamic power occuring due to unproductive flit movement in the network. We also propose a simple logic that precomputes the productive routes of a flit in the succeeding router by adjusting the DWs of the flit.

Concept of Weighted Deflection Count :
In WeDBless, WDC of a flit will be 0 at the time of injection into the network. This WDC is updated at the end of the router pipeline by incrementing or decrementing its value. Frequently deflected flits will have higher WDC value than less deflected ones. Priority is assigned to flits such that flits with high WDC will have higher priority in output port selection.

The four DWs of a flit represent its preference for the four output ports. The DWs can take any one of the three values, -1,+1 or +2 which are coded using 2 bits each. For a flit, the DW of a fully productive output port will be -1 (least weight) whereas DW of partially non-productive ports will be +1 (medium weight). DWs of output ports that deflect flits to the opposite side of destination will be +2 (maximum weight). As an example, if the flit's destination lies in the same column and towards south of the current router, then DWs in the north, south, east and west directions are +2,-1,+1 and +1 respectively. In order to incorporate the WDC and DWs, we use additional 14 bits in the flit header; 6 bits for WDC and 8 bits for DWs.

Router Pipeline:
WeDBless uses a two stage pipeline with one cycle latency each. In the WeDBless architecture as shown in Figure 1,
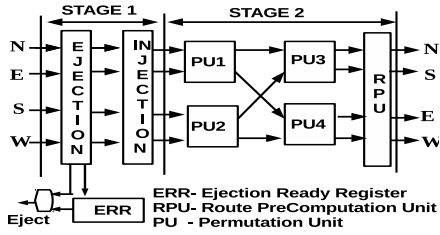
Figure 1: WeDBless router pipeline.

incoming flits enter the router pipeline through input ports (shown on left) and move towards output port (shown on right) through various units. Ejection and Injection constitute the first cycle of the router pipeline which is similar to pipeline architecture of CHIPPER [2]. In the Ejection unit, we provide an Ejection Ready Register(ERR) which buffers one among the multiple flits destined to the local core. This buffering serves to reduce deflections due to such flits roaming around the network.

In the next pipeline stage, flits pass through the permutation network at the end of which output ports are alloted to them. WeDBless router uses the permutation network proposed in CHIPPER for output port selection. In the port allocation stage, a flit competes to occupy the output port with lowest DW. After port allocation, the DW value of allotted output port is added to the WDC of the flit. The incrementing and decrementing of WDC using DWs helps to maintain its value between 0 and 63.

The Route Precomputation Unit (RPU) is placed at the end of the router pipeline and after this unit, flits proceed to the next router through the flit channel. The RPU recalculates the four DWs of the flit for the next router. It consists of simple hardware which can increment, decrement or retain the previous value of DWs of the flit depending on which output port is allocated to it.

3. EXPERIMENTAL METHODOLOGY

We model BLESS and CHIPPER router designs with two cycle latency by modifying the traditional cycle accurate NoC simulator, Booksim [1]. We also model the WeDBless router design by making modifications to the CHIPPER simulation model as mentioned in Section 2. We conduct all evaluations using single flit packets. Using synthetic traffic patterns, we conduct experiments for 8x8 mesh network. After providing sufficient warm up, we collect the deflection rate and average flit latency for various flit injection rates from zero load to saturation.

Results:

Deflection of flits through the network causes unnecessary dissipation of dynamic power. The aim of WeDBless is to minimise these deflections and achieve energy efficiency for the NoC. Average deflection rate is computed as the average number of deflections encountered per flit. From evaluations using synthetic traffic, we observe that WeDBless reduces the deflection rate by a maximum of 56% compared to CHIPPER. The unique routing mechanism in WeDBless reduces deflections in three ways : (1) by computing more than one productive paths for a flit (2) by increasing WDC value for deflected flits so that they win desired output ports in the succeeding router's arbitration (3) by providing an
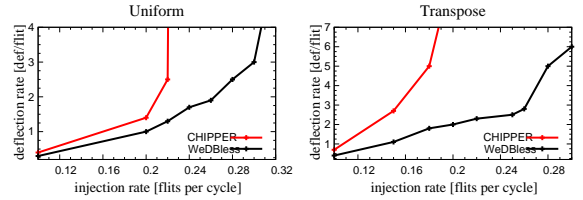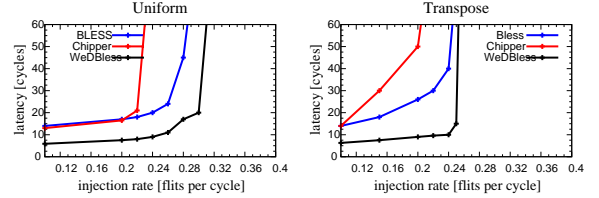

Figure 2: Avg. deflection rate for 8×8 mesh

Figure 3: Avg. flit latency for 8×8 mesh

ERR that buffers one among multiple flits destined to the local core.

Reduction in deflection rate leads to lower average latency as well. For uniform traffic, WeDBless improves the network saturation point by 8% compared to BLESS. The introduction of ERR and the concept of port allocation using WDC are responsible for the reduction in latency for WeDBless compared to BLESS. Compared to CHIPPER, WeDBless exhibits improvement in network saturation point by 55% for transpose traffic and 26% for uniform traffic.

4. CONCLUSION

We proposed a novel bufferless deflection router for mesh NoCs that reduces the deflection rate and average latency of flits in the network. We used Weighted Deflection Count to rank flits and Directional Weights to prioritise output ports for a flit. We also presented a simple and effective method to precompute the routes of a flit by recomputing the directional weights . Future work on WeDBless router consists of evaluating the network using benchmark applications and analysing the application level speed up. Further, power and area modelling of WeDBless router can be conducted to prove the energy savings obtained due to the reduced deflection rate.

5. ACKNOWLEDGMENT

6. REFERENCES

[1] W. Dally and B. Towles. *Principles and Practices of Interconnection Networks*. Morgan Kaufmann, USA, 2004.

[2] C. Fallin, C. Craik, and O. Mutlu. Chipper: A low-complexity bufferless deflection router. In *17th International Symposium on High Performance Computer Architecture*, pages 144–155. IEEE Computer Society, 2011.

[3] T. Moscibroda and O. Mutlu. A case for bufferless routing in on chip networks. In *36th Annual International Symposium on Computer Architecture*, pages 196–207. ACM, 2009.