

ADAPTIVE ADMISSION CONTROL ON THE SPINNAKER MPSOC

Shufan Yang, Steve B. Furber, Luis A. Plana
School of Computer Science, The University of Manchester
Oxford Road, Manchester M13 9PL
Email: shufan.yang, sfurber, plana @cs.man.ac.uk

Abstract—End-to-end communication service is critical to maximise both flexibility and performance on a Multi-Processor System-on-Chip (MPSoC). We introduce adaptive admission control to ensure fair bandwidth allocation to each processing node on an MPSoC platform. The results from the Matlab system model show good agreement with the experimental results from the HDL model. Consequently, the Matlab model can be used as an effective prototyping toolkit.

I. INTRODUCTION

Multi-Processor Systems-on-Chip have strict performance, power and area cost goals. A key requirement to enable MPSoC platforms to handle real-time, high-demand applications is on-chip communication service guarantees. Many MPSoC applications are computing intensive and require very complex coherence algorithms for efficient parallel computation. An unpredictable end-to-end communication latency is unacceptable. Although there have been several studies of service guarantee techniques for on-chip communication, many challenges remain.

The SpiNNaker project is developing a new architecture that uses lessons from neuroscience to inspire an innovative approach to massively-parallel computer design [1]. The architecture is primarily aimed at simulating a billion spiking neurons in real time, but also offers a novel high-performance computing solution.

Fig. 1 shows a simplified block diagram of the SpiNNaker chip. Each processor node is a complete processing sub-system with local memory and DMA capability. It is connected to its environment (which includes off-chip SDRAM memory and other on-chip components for system configuration) using the CHAIN asynchronous interconnect [2]. Although CHAIN provides a highly-reliable and cost-effective fabric, it is a best-effort interconnect, and as such cannot guarantee to meet the bounded-service requirements of communications between the proces-

sors and the off-chip SDRAM. However, the bandwidth available at the SDRAM interface must be fairly shared to balance the distributed task of transferring data to and from this memory.

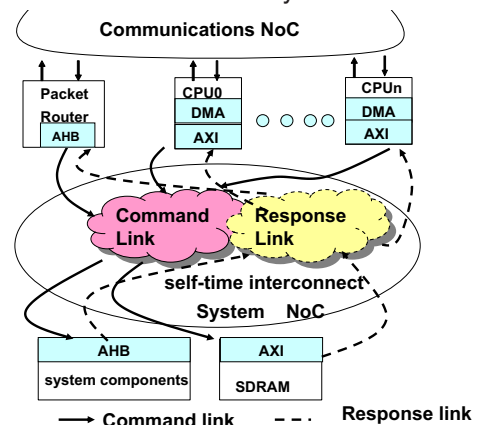


Fig. 1. The SpiNNaker chip block diagram

In this paper we introduce an adaptive admission control mechanism to satisfy the communication demands of applications running on the SpiNNaker chip. The proposed adaptive admission control maintains fairness in the face of the unbalanced bandwidth allocation characteristics of the fabric through the use of a closed-loop feedback system. By properly optimising the parameters we are able to obtain an approach that can dynamically converge to a fair bandwidth allocation and prevent potential network saturation problem.

II. PROBLEM STATEMENT

Fig. 2 shows a simple network topology with three initiators (processors) accessing one target (a shared memory). The on-chip interconnect has two dedicated communication links: the command link is used by the initiators to initiate a communication transaction, and the response link is used by the targets, such as the SDRAM controller, to respond to transaction requests.

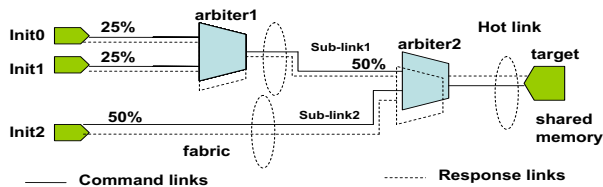


Fig. 2. Low-complexity interconnect example

In steady state, a sequence of burst read requests may create congestion in the command link. Ultimately the 2-way asynchronous arbiters, which merge the requests from the initiators into a sequenced stream, will transfer the congestion back-pressure to all incoming links and the fabric will saturate. In this condition the interconnect will behave unfairly, as a direct result of the binary tree arbitration structure. If all three initiators request continuously, they will not be served equally. Under continuous requests from both sides, asynchronous arbiters will alternate their grants. Here *arbiter2* grants 50% of the bandwidth to *sub-link1* and 50% to *sub-link2*, as shown in Fig. 2. Similarly, *arbiter1* will grant half of the *sub-link1* bandwidth to *initiator0* and half to *initiator1*. Over a period of time, therefore, the bandwidth allocation to each of *initiator0* and *initiator1* is 25% of the total, whereas that allocated to *initiator2* is 50% of the total, leading to a natural unbalance of the system towards *initiator2*.

Note that the bandwidth allocation varies with the size of the buffer in the target. If the target has an 8-command FIFO, the system delivers a mildly asymmetric bandwidth allocation. This is because the target FIFO increases the capacity of the fabric to absorb commands and thereby reduce congestion. As the network capacity is increased, the competition among initiators in the arbitration tree is reduced. Increasing the target buffer size sufficiently can solve the fairness problem altogether, but in practice the large buffers can be too expensive.

III. ADAPTIVE ADMISSION CONTROL

As we have seen, as long as the fabric is not operating near its saturation point, back-pressure will not arise in the command fabric and the network will deliver a fair bandwidth allocation.

In this paper, we propose an adaptive admission control mechanism to regulate the traffic injection rates. The adaptive admission control works as a local scheduler by sensing global system latency. The local admission control regulates the packet injection rate at the source to keep the latency below a threshold. The predefined threshold is the mean system latency when all initiators have an equal

share of the full bandwidth of the memory interface. If some initiators are not using their full bandwidth share, the system latency will reduce and other busy initiators can increase their demands on the fabric to exploit the additional capacity available to them.

The system is effective because there is a region of operation where the target bandwidth is fully utilized but the command fabric is unsaturated and the arbitration therefore fair. The role of the admission control mechanism is to hold the system within this region whatever the number of active initiators.

A. System model

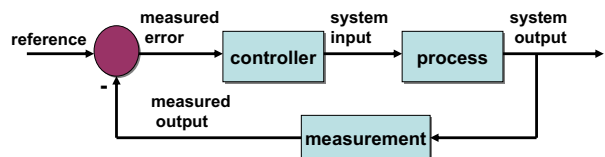


Fig. 3. A basic admission control model

The admission control mechanism is a conventional feedback control system as illustrated in Fig. 3. The system is a set of units working together to deal with the dynamics of the network-on-chip. The desired set-point is called the reference, which here is the system latency at the threshold. The output of the system is the packet injection rate control. Each local admission controller measures the latency of local transactions to estimate the load on the network fabric.

Fig. 3 shows a negative feedback loop, which means that the measured output is subtracted from the reference to create the error signal that is amplified by the controller. Then, the controller uses the error signal to modify the process in order to eliminate the error. This closed-loop control system has the merit of being able to match the measured values to the required values. However, a problem can arise if there are delays in the system. Such delays cause the corrective action to be taken too late and can lead to oscillation and instability. To accommodate the delay in the feedback loop, a filter is used to reduce high-frequency noise and to help the system settled into a steady state.

B. MATLAB results and discussion

We use Matlab to construct a discrete-time model. By modeling the closed-loop admission control system, we can determine the effect of different gains at an early stage in the design process.

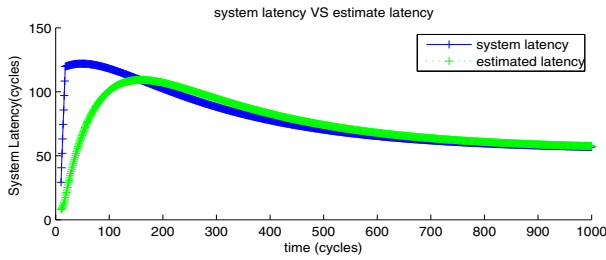


Fig. 4. Plot of System latency VS estimated latency with the proportional control gain $K = 0.001$

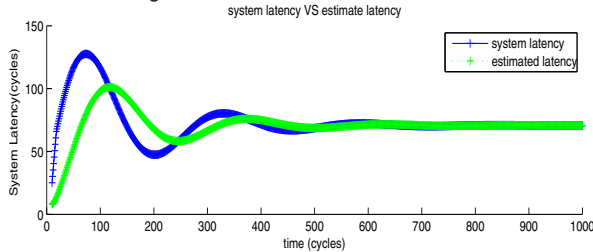


Fig. 5. Plot of System latency VS estimated latency with the proportional control gain $K = 0.01$

The admission control system is an independent proportional control system [3]. A proportional controller uses the filtered error signal multiplied by a constant to adjust the process. The resulting simulated dynamics of the system latency with two different proportional control gains are shown in Fig. 4 and Fig. 5. It is clear that the proportional gain affects the speed of the response: a larger proportional gain improves the response time of the system, the steady-state error decreases and the overshoot increases. The Matlab model allows the effect of different gains to be determined. For the 5-initiator-to-1-target case, a proportional gain around 0.01 gives a good balance between responsiveness and stability.

Stability of response is required for any combination of control system conditions. Suppose initially the system latency is smaller than the reference point. The controller then increases the traffic input rate. Because of the time delay, some time will elapse before the local admission control senses any change. The result will be that the controller will continue to increase the input data rate and the system latency will draw past the reference point. Then, the controller will decrease the traffic input rate according to the measured error. Again because of the time delay, before the local admission control senses the system latency will draw under the reference point. And the controller will increase the traffic inputs rates. This process is described by conventional control theory, and depending on the time delays and loop gain the system may be over-damped, critically-damped, under-damped or

unstable.

IV. EVALUATION AND DISCUSSION

In this section we evaluate the adaptive admission control strategy at circuit level. We have developed structural Verilog code for the asynchronous interconnect using the CHAINworks tool [5] and written RTL code for the admission controller.

A. Evaluation platform

The experimental scenarios we have explored are: five initiator devices using the AXI [4] protocol connected to one target device using the AXI protocol. In theory, the initiators can issue requests as often as they require. However, the number of outstanding commands an initiator can issue is constrained by the capability of the interconnect interface. Currently the design only supports 8 outstanding commands, which means that at most 8 read commands can be issued before the corresponding data returns. The traffic profiles are built manually using a fixed 8-word burst for each transaction.

B. Practical implementation

A successful hardware implementation of the adaptive admission control must take account of performance criteria and have a low area overhead. Typically, proportional control is implemented with floating-point hardware, which simplifies the program but is expensive to implement. It is possible, however, to implement proportional control with fixed-point hardware by choosing byte fractional formatting that eliminates some calculations.

In addition, both finite impulse response (FIR) and infinite impulse response (IIR) filters are candidates for the filter in the feedback loop. Of these the FIR implementation require more memory to achieve a given filter response characteristic.

Fig. 6 shows the function blocks in the adaptive admission control design. Each input port is the asynchronous handshake signals with the local IP block. The shift register is for sampling the requests. The delay counter keeps the last time delay record. The controller model calculates the predicted time interval of grants according to the measured error. Then the comparator enables a grant if the current request interval is bigger than the computed value.

A better performance guarantee can be attained at the expense of increased silicon area. The estimated area is $0.02mm^2$ based on UMC130nm technology. One fixed-point multiplier, an adder, two 32-bit registers and 32x32-bit ROM contribute to the total area.

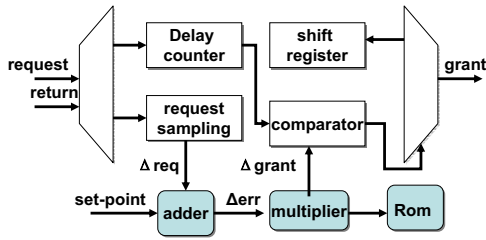


Fig. 6. A datapath block diagram of admission controller

This leads us to conclude that distributed adaptive admission control will be efficient and flexible for larger-scale system.

C. Comparison between Matlab model and verilog model

We must ensure that our model is accurate for the range of input conditions under which the system will operate. We acquire experimental data from a 5-initiator-to-1-target interconnect experiment over heavy traffic load. To account for potential differences across the simulation, we chose 16 monitor points and compared our test data with results from simulation on the Matlab model under the same input rates. The differences between these values and the Verilog experimental results are calculated. The criterion of model accuracy is mean relative error. The mean relative error is below 0.09 and gives a 90% confidence of accuracy in the steady state, so it is acceptable for our application.

D. Evaluation

Scheme types	Ave. system latency	Max. system latency
With AC	232ns	299ns
Without AC	323ns	570ns

TABLE I
LATENCY COMPARISON

Table I shows that the adaptive admission control approach offers many possibilities for performance improvement. When the offered load is the maximum achievable throughput, the average and maximum latency in the uncontrolled fabric are found to be 323ns and 570ns respectively. For the system with admission control, the average transaction latency becomes 232ns. This reduction in latency is mainly due to the proposed admission control regulating the data rates into the fabric directly. As a result, the maximum of latency in the system with admission control drops from 570ns to 299ns, which is close to a 50% reduction.

Input	Without AC	Percent.	With AC	Percent.
Init0	800	25%	764	19%
Init1	800	25%	732	18%
Init2	401	12%	778	20%
Init3	402	13%	800	22%
Init4	799	25%	799	21%

TABLE II
SIMULATION RESULTS

In our Verilog simulation we estimate bandwidth allocation by measuring the total number of transactions completed by each initiator over a given period of time. In table II, the second column presents the results from a system without admission control (AC). This shows the number of transactions completed by each initiator and indicates the unfair allocation of bandwidth. The fourth column shows the number of transactions completed by each initiator in the same system with our admission control. The system with admission control reduces bandwidth allocation unfairness to low level.

V. CONCLUSION AND FUTURE WORK

We propose a distributed adaptive admission control mechanism based on closed-loop feedback. The admission decision can be made in real time and without much computational effort, and the admission control works effectively.

In future work, we will extend our adaptive admission control algorithm to general network-on-chip interconnect and consider the impact of multiple targets on our algorithm.

ACKNOWLEDGMENT

S. Yang would gratefully like to acknowledge the support from the national 863 Program (2007AA012104), the School of Computer & Communication, Hunan University, China. The authors would like to acknowledge the support for this work from the Engineering and Physical Sciences Research Council through EPSRC grants GR/S61270/01 and EP/D07908X/1.

REFERENCES

- [1] Steve Furber, Steve Temple and Andrew Brown, "On-Chip and Inter-Chip networks for Modelling Large-Scale Neural Systems", *International Symposium on Circuits and Systems, ISCAS*, pages 21-24, May 2006.
- [2] John Bainbridge and Steve Furber, "Chain: a Delay-insensitive chip area interconnect," *IEEE Micro*, volume 22, pages 16-23, Sept. 2004.
- [3] Martin Healey, "Principles of Automatic Control", *The English universities press LTD*, pages 45-78, 1976
- [4] ARM Ltd, "AMBA Specification (ARM IHI 0022B)", www.arm.com, 2007
- [5] Silistix Ltd., "Building and Analyzing On-Chip Networks using CHAIN architect", *Technical report*, V1.2.5, AUG 2008