

單晶片多處理機可程式實驗平台之設計與實現(II)

Design and Implementation of a Programmable Experimental Platform

for a Single Chip with Multiple CPUs (II)

計劃編號: NSC87-2213-E-009-030

執行日期: 86/8/1 - 87/7/30

主持人: 王國禎 國立交通大學資訊科學系副教授

一、中文摘要 (關鍵詞: 現場可程式邏輯陣列閘、多處理機、單邊縱橫交換鍵、疊流式、產出量、硬體描述語言。)

在本計劃中, 我們提出了一個建構在單晶片多處理機可程式實驗平台的疊流式單邊縱橫交換鍵。疊流式單邊縱橫交換鍵是由單邊縱橫交換鍵、單邊縱橫仲裁器、以及疊流式介面三大部分所組成。疊流式介面控制了疊流式協定中仲裁、請求、窺探、回應和資料共五個階段的進行。它是我們提出的架構中最主要的部分。單邊縱橫交換鍵接收來自處理機之疊流式介面的匯流排配置請求, 以控制單邊縱橫交換鍵來達成處理機和記憶體模組之間的連接。我們比較了疊流式、傳統式、改良式和漣漪式四種單邊縱橫交換鍵的產出量。實驗數據顯示 1×1 疊流式單邊縱橫交換鍵的最大產出量 (其對應的等待時間) 是其它三種交換鍵的 4.1 倍 (0.55 倍)。在隨機故障模式下, 我們也估算出四種交換鍵的等待時間。此外, 我們使用 Verilog 硬體描述語言及 Xilinx 的 FPGA (現場可程式邏輯陣列閘) 分別描述及實現了一個 4×4 疊流式單邊縱橫交換鍵。Verilog 模擬的結果驗證了我們的疊流式設計功能。疊流式單邊縱橫交換鍵的成本效益 (產出量/面積) 是傳統式的 4.5 倍及改良式或漣漪式的 4.1 倍。本研究的貢獻是設計出一個高產量及具成本效益的疊流式單邊縱橫交換鍵以配合高效能的多處理機, 從而消除效能瓶頸。

英文摘要 (**Keywords:** FPGA, pipelined, multiprocessor, one-sided crossbar switch, throughput, Verilog)

In this project, we propose a pipelined one-sided crossbar switch for a single-chip multiprocessor system. A pipelined interface for each processor is the main part of our proposed structure, which controls the five phases of our pipelined protocol: *arbitration, request, snoop, response, and data*. We compare the throughput of the four one-sided crossbar switches: the pipelined, traditional, modified, and ripple K one-sided crossbar switches. Experimental results show that the maximum throughput (the corresponding waiting time) of the 8×8 pipelined one-sided crossbar switch is about 4.8 times higher (0.55 times lower) than those of the other three one-sided crossbar switches. We also evaluate the waiting time of these four switches under random faults. In addition, we have described and implemented a 4×4 pipelined one-sided crossbar switch using Verilog HDL and Xilinx FPGAs, respectively. The cost-effectiveness (in terms of throughput/area) of our pipelined one-sided crossbar switch is 3.5 times higher than that of the traditional one-sided crossbar switch and 2.8 times higher than that of the other two switches. The contribution of this work is designing a high throughput and cost-effective pipelined one-sided crossbar switch

to match high performance multiprocessors and to eliminate performance bottleneck.

二、計劃緣由與目的

內接網路是一個高效能多處理機系統中極重要的部分。雙邊縱橫交換鍵具有不阻塞的高頻寬特性，但是不具有容錯性。單邊縱橫交換鍵改良了雙邊縱橫交換鍵之缺點，具有容錯及不阻塞的特性，極適合用來作多處理機之內接網路。在傳統式單邊縱橫交換鍵中，處理機和記憶體模組是放置在同一邊，而它們之間的資料和訊息傳遞則是透過交叉節點來連結。由於傳統式單邊縱橫交換鍵的成本過高，所以有人提出了改良式和漣漪式單邊縱橫交換鍵。這兩種新式的單邊縱橫交換鍵減少一些交叉節點而仍保持著一定的容錯能力。

匯流排競爭及記憶體衝突是影響內接網路表現之兩項主要因素。為了降低匯流排競爭可利用不阻塞之高頻寬內接網路。為了減少記憶體衝突則可增加記憶體模組個數或縮短記憶體存取時間。另外，疊流式控制也能降低上述因素所造成的效能影響。疊流式控制是當某筆交易等待資料被送回時，釋放匯流排的使用權供其它的交易使用。傳統的方式則需等待資料送回才釋放匯流排。利用疊流方式可使內接網路的產量提高。現今許多高效能的微處理機，像是 Pentium Pro 和 Power PC，為了達成計算的高平行性，都提供了疊流式的匯流排協定來處理記憶體和輸出入裝置的存取。

在本篇報告中，我們致力於高效能疊流式內接網路之設計，以降低記憶體衝突所造成的影響。當某匯流排因故障而無法使用時，疊流式設計也能降低匯流排競爭的發生機會。我們以硬體描述語言來描述疊流式單邊縱橫交換鍵及以模擬來驗證內接網路的功能正確性。此外，我們使用

FPGA 合成軟體來分析延遲時間及面積，以進一步評估成本效益。

三、研究方法及成果

圖(一)顯示疊流式介面及其所連結的單邊縱橫仲裁器與單邊縱橫交換鍵，分述如下：

- 單邊縱橫交換鍵：它是一個不阻塞的交換鍵，而且每條匯流排包括一個位址匯流排和一個資料匯流排。它們的寬度都是32位元。
- 單邊縱橫仲裁器：它決定哪一個請求將被配置匯流排，以建立連接通路。
- 疊流式介面：它從處理機接收位址（讀取交易）或位址和資料（寫入交易），同時它也以疊流的方式來傳送位址與資料。

在疊流式介面中有三種型式的緩衝區，它們分別是：位址緩衝區、送出資料緩衝區和接收資料緩衝區。而每個交易在請求和資料階段分別會送出匯流排配置的要求，所以每個交易必須經過兩次的匯流排配置。我們提出的疊流式協定有下列五個階段：

- 仲裁：處理機以循環的方式將一個交易選出，然後相關位址（和資料）接著被送至疊流式介面。
- 請求：疊流式介面送出位址給單邊縱橫仲裁器來配置匯流排。如果沒有記憶體衝突的問題，則交易進入下一個階段；否則就須進行重送位址的動作。
- 窺探：在這個階段，某交易已完成了請求，而這時需要等待記憶體模組的就緒。
- 回應：目的裝置（這裏指記憶體模組）送出回應給提出請求的疊流式介面，這表示記憶體模組或處理機可以送出資料了。
- 資料：如果沒有處理機的衝突，則資料可以傳送至匯流排上。

此外，有一個錯誤偵測器，可以在請

求階段偵測位址錯誤以及在資料階段偵測資料錯誤。如果有錯誤發生，位址或資料將會被重送。圖（二）顯示出一條疊流式匯流排上有四個交易的例子。四個交易分別是：R1、R2、W1 及 W2，其中 R1、R2 是讀取交易而 W1、W2 是寫入交易。在窺探階段，R1 花了三個時脈週期，而 R2 花了四個時脈週期去等待記憶體模組就緒。這個疊流式的協定允許多個交易共存在窺探階段。在 W2 的請求階段，可能發生了位址錯誤或記憶體衝突，所以就必須進行重送。

我們使用 Verilog HDL 來設計疊流式單邊縱橫交換鍵，其樹狀分解圖如圖（三），其樹葉部分的模組描述如下：

- 選擇器：選出一個匯流排來傳送位址或資料。
- 移位器：透過移位器不斷的輪替交易優先權，以避免飢餓的情形發生。
- 交叉節點陣列：包含疊流式單邊縱橫交換鍵的所有交叉節點。
- 解多工器：主要是將匯流排配置的信號轉換為交叉節點的起動信號。
- 錯誤偵測器：檢查位址或資料是否有錯誤發生。
- 緩衝區：分為位址、送出資料和接收資料緩衝區，主要是暫存位址或資料。
- 緩衝區控制器：它們分別管理疊流式介面中的三個緩衝區。

在成功地模擬四種單邊縱橫交換鍵之後，我們使用 Synopsys 的 FPGA 編譯器來估算出四種交換鍵的面積和延遲時間。在表（一）中，我們可以看出疊流式交換鍵的面積和延遲時間比傳統式交換鍵分別增加 38% 和 21%。

四、討論與結論

我們比較了疊流式和其它三種現存的單邊縱橫交換鍵之產出量和等待時間。因為在疊流式單邊縱橫交換鍵中，每個交易必須進行五個階段，我們假設基本時間單

位是一個時脈週期。我們採用了最大流量演算法中的 Lift-to-Front 演算法，來模擬四種單邊縱橫交換鍵的產出量。圖（四）顯示在隨機故障模式下，四種 8x8 單邊縱橫交換鍵最大的產出量。有一項研究顯示交換鍵的產量為 99%，那就是說，對一個 8x8 的交換鍵來說，交叉節點的平均故障數是少於二個的。所以我們的疊流式單邊縱橫交換鍵是可行的，並且可維持高產出量。圖（五）顯示在隨機故障模式及最大產出量下的等待時間。實驗數據顯示，1x1 疊流式單邊縱橫交換鍵的最大產出量（其對應的等待時間）是其它三種交換鍵的 4.1 倍（0.55 倍）。在圖（六）中，疊流式單邊縱橫交換鍵的成本效益（產出量/面積）是傳統式的 A.5 倍及改良式或漣漪式的 A.1 倍。

疊流式和其它三種單邊縱橫交換鍵最主要的不同在於：每個處理機有一個疊流式介面。即使一個疊流式的交易需要做二次的匯流排配置，我們發現多出來的那一次配置對效能的影響是十分輕微的。因此，我們的疊流式單邊縱橫交換鍵確實具有高效能的特性，可以解決常發生在多處理機系統中內接網路效能瓶頸的問題。

五、參考文獻

- [1] K. Hwang, "Advance Computer Architecture: Parallelism, Scalability, Programmability," *McGraw-Hill*, 1993.
- [2] F. Pong, M. Browne, A. Nowatzyk, and M. Dubois, "Design Verification of the S3.mp Cache-Coherent Shared-Memory System," *IEEE Trans. on Computers*, pp. 135-140, Jan. 1998.
- [3] J. Ghosh, A. Varma, and Krishnamurthy, "Distributed Control Schemes for Fast Arbitration in Large Crossbar Networks," *IEEE Trans. on VLSI Systems*, pp. 54-67, Mar. 1994.
- [4] K. Wang and C. K. Wu, "Design and Simulation of Fault-Tolerant Crossbar

Switches for Multiprocessors Systems," *Proceedings of Pacific Rim International Symp. Fault Tolerant Systems*, pp. 198-203, Dec 1995.

- [5] N. Sarangdhar and G. Singh, "An Overview of the Pentium Pro Processor Bus," *Digest of Papers, COMPCON '96*, pp. 383-387, Feb. 1996.
- [6] M. S. Allen, W. K. Lewchuk, and J. D. Coddington "A High Performance Bus and Cache Controller for PowerPC Multiprocessing Systems," *Proceedings of the ICCD '95*, pp. 204-211, Oct. 1995.
- [7] S. Brown and J. Rose, "FPGA and CPLD Architectures: a Tutorial," *IEEE Design and Test of Computers*, pp. 42-57, June 1996.
- [8] T. H. Cormen, C. E. Leiserson, and R. L. Rivest, "Introduction to Algorithms," *McGraw-Hill*, 1990.
- [9] K. S. Hedlund and L. Snyder, "Systolic Architectures - A Wafer Scale Approach," *IEEE Int. Conf. Computer Design: VLSI in Computers*, pp. 604-610, Oct. 1984.

圖 (二) 疊流交易之時序圖。

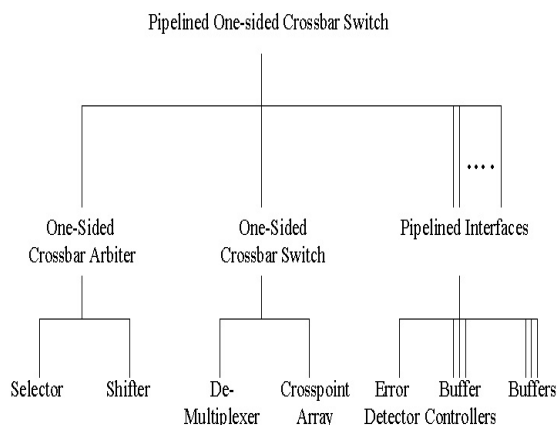
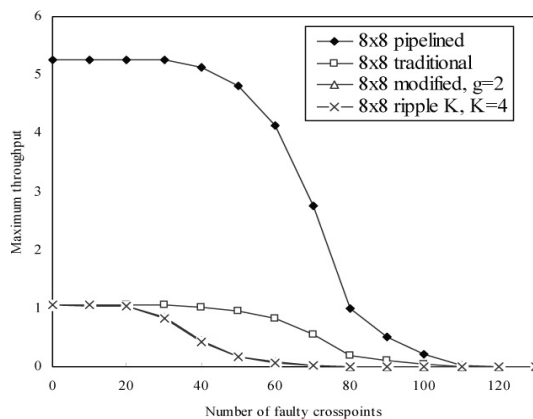
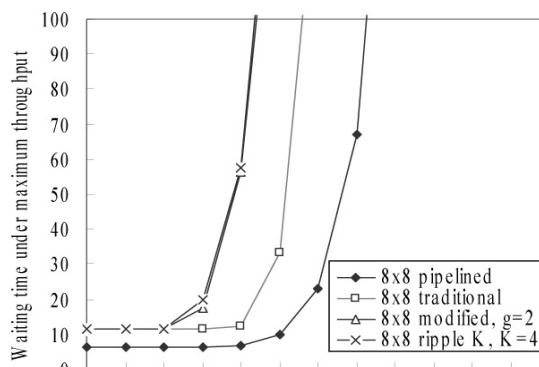


圖 (三) 單邊縱橫交換鍵設計之樹狀分解圖。圖 (四) 隨機故障模式下，四種8x8單



邊級橫交換鍵之最大產出量。

圖 (五) 在隨機故障模式及最大產出量下，四種8x8單邊級橫交換鍵的等待時間。



六、圖表

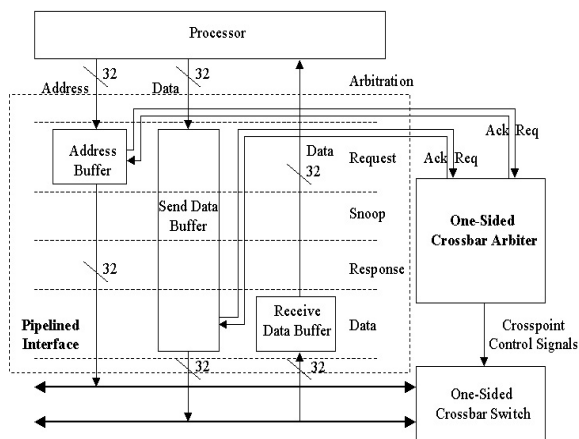
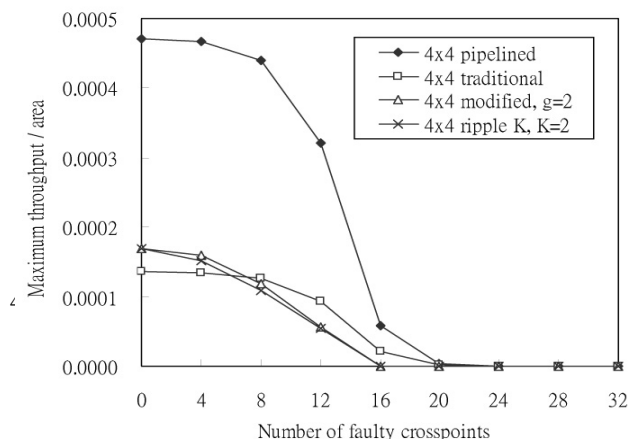


圖 (一) 疊流式介面之方塊圖。

Clock	[Clock waveform]															
Arbitration	R1	R2		W1		W2										
Request		R1	R2		W1		W2	W2								
Snoop			R1	R1	R1	R2	R2	W1	W2							
Response						R1		R2	W1							
Data								R1		R2						
Address Bus	[Address bus waveform]															
Data Bus	[Data bus waveform]															



圖（六）隨機故障模式下，四種4x4單邊級橫交換鍵之成本效益。

表（一）面積和延遲時間的報告。

Switch	Pipelined	Traditional	Modified	Ripple K
Total Area	5577	4033	3233	3233
Total Delay	116.1ns	95.7ns	92.9ns	92.9ns