

# CAD/VLSI Circuit Design 期末報告

## 構想說明

7109064300 范文軒 電機丁組(數位)

### (一)、動機

搜尋演算法是近些年科技中很重要的基石，而字串比對演算法(string matching algorithm)也是其中之一。像這樣的比對方式會被使用至 DNA 序列的比對、資料庫的查詢(query)等。而如今資料庫系統中的資料量越來越大。為了某個特徵比對，可能會使用到多次的比對查詢。因此加速這個過程，就顯得重要的。若使用暴力破解的方式，必須一個字元依次比較，而因為沒有辦法加速，時間複雜度會為  $O(m*n)$  (in worst case)。如今不同新的演算法也不斷推出，像是近年最多人使用的演算法之一：KMP 演算法<sup>[1]</sup>，優化了最原始的方式，使得無意義的比較可以跳過，帶來更多速度上的優勢，而時間複雜度減至  $O(m+n)$  (in worst case)。而本次構想希望參考可平行化運算 KMP 演算法<sup>[2]</sup>，並且利用硬體亦可平行化的優勢，實做出有平行化的 KMP 演算法的 SME(string matching engine)。

### (二)、主要概念與設計<sup>[2]</sup>

#### (1). KMP Algorithm<sup>[1]</sup>

KMP 演算法相較於暴力破解，多增加了前處理(preprocessing)的部分，在此演算法即所謂的 Failure Function。增加了一個關於 Pattern 的表格去紀錄若判別失敗該跳去哪一個位置。

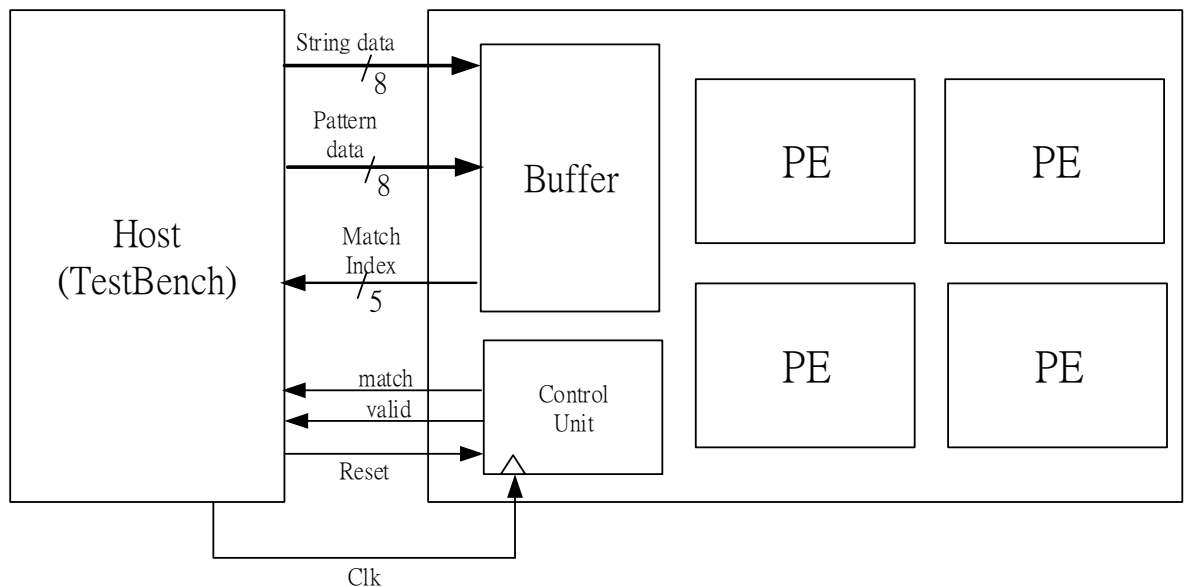
#### (2). 可平行化 KMP Algorithm<sup>[2]</sup>

而在參考[2]中，使用了平行化技巧，將比對時分切成 4 個特殊大小的字串同時比對，進而減少時間的損耗。

參考演算法概念後，在構想設計中設計以下狀態：Idle、String、Pattern 輸入、Failure Function 建立、比對環節以及輸出結果。並打算在比對環節中設計 4 的 PE(Process Engine)同時運作於比對環節中進而實驗平行化，並將以上的概念組合成有限狀態機實現 SME。

### (三)、主要架構

預計完成 ASCII 編碼的 SME。並且設定輸入字串上限 32Bytes，且 Pattern 最大為字串長度上限為 8Byte。



圖一、SME 系統架構圖

### (四)、規格

Hardware	
Synthesis Process	TSMC 90 nm
CLK Cycle	100Mhz

### (五)、參考

- [1] D. E. Knuth, J. H. Morris and V. R. Pratt, "Fast Pattern Matching in Strings," *SIAM J. COMPUT.*, Vol.6, No.2, June 1977.
- [2] U. S. Alzoabi, N. M. Alosaimi, A. S. Bedaiwi and A. M. Alabdullatif, "Parallelization of KMP String Matching Algorithm," *2013 World Congress on Computer and Information Technology (WCCIT)*, pp. 1-3, 2013.