設計內容

[1] 設計者姓名與連絡電話

學生姓名: 鞠敦揚、廖昶翔、施伯儒

聯絡電話: 0934224900(鞠敦揚)

Email: b09901188@ntu.edu.tw b09901178@ntu.edu.tw b09901081@ntu.edu.tw

[2] 專題名稱

中文專題名稱:硬體加速高斯-賽德爾迭帶法求聯立方程式及反矩陣

英文專題名稱:Solving Simultaneous Equations and Inverse Matrices on Gauss-Seidel Iteration Machine Acceleration Chip

[3] 全新設計或改版說明

此案件為設計者全新設計。

[4] 原理及架構說明

高斯-賽德爾迭帶(Gauss— $Seidel\ method$)是數值線性代數中的一個迭帶法,可用來求出形如 Ax = b 的線性方程組解的近似值。該方法以卡爾·弗里德里希·高斯和路德維希·賽德爾命名。

對於一個 Ax = b 的線性方程組,其中 A 是一個 $n \times n$ 的矩陣,x 和 b 是 n 維向量。高斯-賽德爾迭代法的求解過程如下:

- 1. 首先,將x的初始值設置為一個任意的向量,例如全為0的向量。
- 2. 將 x 的每個分量依次更新為

$$x_i^{k+1} = \frac{1}{a_{ii}} \left[b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{k+1} - \sum_{j=i+1}^{N} a_{ij} x_j^k \right]$$

3. 重複步驟 2, 直到 x 的值收斂到所需的精度或達到一定的迭代次數 k。

高斯-賽德爾迭帶法的關鍵在於每次更新 x 的分量時,使用已經計算出來的新值, 而不是使用上一次的值。這樣可以利用迭帶過程中最新的信息來加快收斂速度。

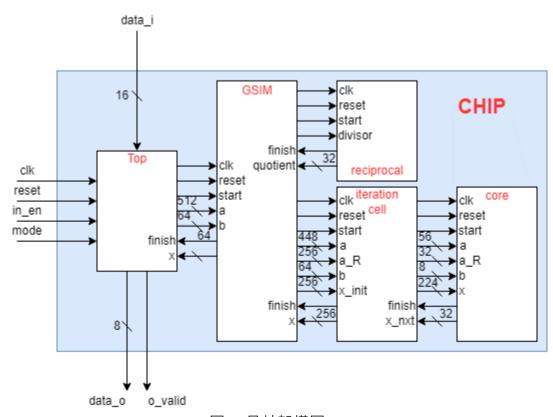
此外關於反矩陣的求解,高斯-賽德爾迭帶法可以應用於求解反矩陣。對於一個 $n \times n$ 的矩陣 A,如果 A 可逆,則其反矩陣 A $A^{-1} = I$ 可以通過求解 n 個線性方程組 Ax[i]

= e[i] (其中 e[i]是 n 維單位向量的第 i 個分量)得到。使用高斯 - 賽德爾迭帶法求解反 矩陣的過程如下:

- 1. 初始化反矩陣的初始值為一個全為 0 的矩陣。
- 2. 對於反矩陣的每一列,分別將初始值的對應元素設置為1。
- 3. 對於每一列,使用高斯-賽德爾迭帶法求解線性方程組 Ax[i] = e[i],其中 x[i]就 是反矩陣的第i列。
- 4. 重複步驟 3, 直到所有列的值收斂到所需的精度或達到一定的迭代次數 k。
- 5. 最終得到的矩陣即為原矩陣 A 的反矩陣 A^{-1} 。

需要注意的是·高斯-賽德爾迭帶法並不適用於所有的線性方程組或矩陣·矩陣 A 需滿足對稱正定矩陣(symmetric positive definite matrix)或對角優勢矩陣(diagonally dominant matrix)。

而此晶片中使用硬體輸入輸出的部份,在解 Ax = b 的線性方程組或解 A 的反矩陣 A^{-1} 中,A 為 8x8 的矩陣、b 和 x 為 8 維向量,輸入的 A、b 為 8 bit 的有號整數 (signed integer),輸出的 x、 A^{-1} 為 24 bit 的有號定點數 (小數點後 16 位)。在初始化 x 時使用的值為 x[i]=b[i]/a[i][i],迭代次數為 8 次。



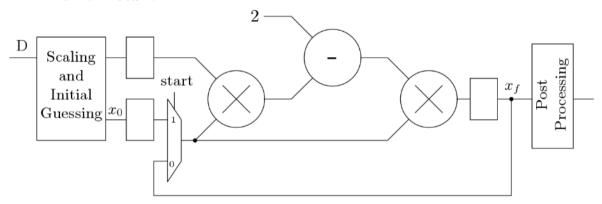
圖一 晶片架構圖

在我們的晶片當中,輸入訊號與輸出訊號是經由 Top 模組進行處理,輸入訊號為一個週期輸入 16 個 bits,經過 32 個 cycle 後 Top 模組會接收完所有的 a 值,再來根據我們輸入的模式決定 b 的值是由外界輸入(模式 0)或是由單位矩陣拆成的向量(模式 1)。

在接收完所有輸入後,會開始進行 GSIM 所需的 x 值的初始化,我們會將 x_i 初始化成 $\frac{b_i}{a_{ii}}$,這部分會使用特殊的除法器,並經過 8 次的處理得出 x 的初始值。

再來將會進行高斯-賽德爾迭帶的運算,運算方式為使用一個計數器來決定每次運算中要輸入進 Core 的 a, b, x 並在進行完運算後,更新 x 並接續進行下一次的運算,每運算八次視為一次迭帶(每個 x 都被更新過一次),而進行完 8 次迭帶後根據模式來決定接下來是進行輸出(模式 0)或是進行下一筆 b 向量的運算(模式 1),而模式 1 會再進行完 8 個 b 向量的運算後進行輸出。

輸出的部分是一個週期輸出 8 個 bits,根據模式來決定輸出的 bits 數:模式 0 為輸出 192bits,模式 1 為輸出 1536bits。



圖二 除法器架構圖

接下來說明除法器部份,採用的是牛頓-拉弗森倒數法(Newton-Raphson Reciprocal),是一種用於求解數字的倒數(reciprocal)的近似值的數值計算方法。

牛頓-拉弗森倒數法的基本思想是從倒數的近似值開始,通過不斷迭帶來改進這個 近似值,直到達到所需的精度。具體的步驟如下:

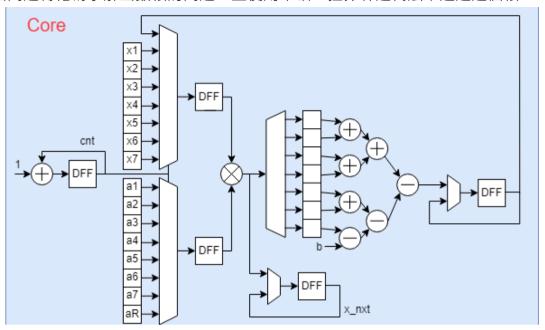
- 1. 首先,選擇一個初始的倒數近似值,這裡我們選擇的是2^{-[log₂ d]},d是要取倒數的數字。
- 2. 使用以下迭代公式來更新近似值:

$$x[n+1] = x[n] * (2 - d * x[n])$$

其中,x[n]是第n次迭代的倒數近似值,d是原始數字。

3. 重複步驟 2 · 直到近似值收斂到所需的精度或達到一定的迭代次數 k · 這裡選擇的 k 為 12 次。

在每次迭代中,通過使用迭代公式,倒數的近似值被不斷改進。該方法的關鍵在於 將原始問題轉化為求解函數根的問題,並使用牛頓,拉弗森迭代法來逼近這個根。



圖三 運算核心架構圖

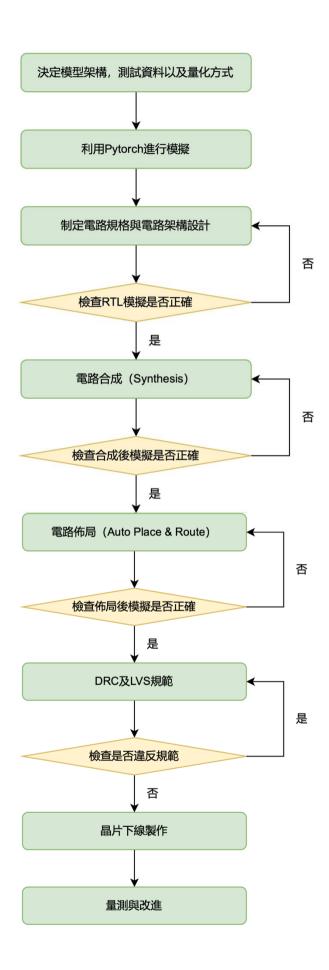
再來是運算核心的說明,首先會有7個周期來將輸入的 x_j 與 a_j 乘起來並存到一個7*32的矩陣裡面,經過7個週期後會將 b_i 與這個矩陣裡的7個值相減,並在下個週期乘上 a_{ii} 的倒數後進行輸出,此為一次運算的過程,在我們的晶片中,運算核心會重複進行此運算8*8次來得到一組x的近似解。

經過此晶片計算得出的結果 x (模式 0) 或 A^{-1} (模式 1) 與 A 重新計算後得出新的 b (模式 0) 或 I (模式 1),與給定的 b 或 I 進行比對後,計算 total square error 的平均值約為 10^{-7} 。

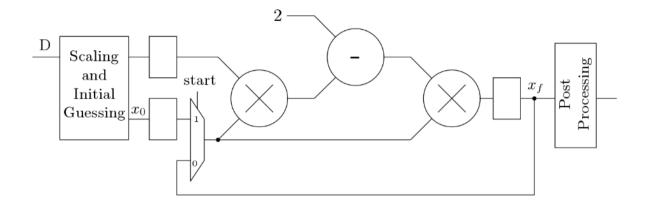
Signal Name	I/O	Width		
clk	I	1		
reset	I	1		
in_en	I	1		
mode	I	1		
data_i	I	16		
out_valid	О	1		
data_o	O	8		
Total: 29 signal				

表一、輸入及輸出腳位

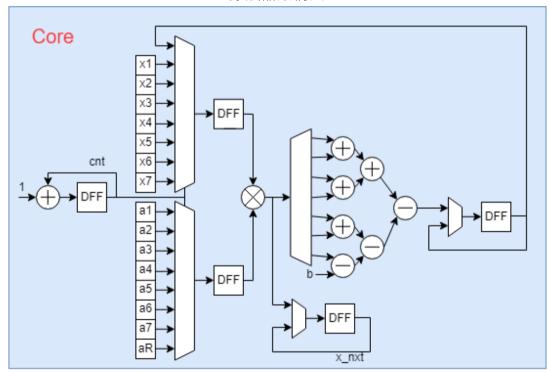
[5] 設計流程



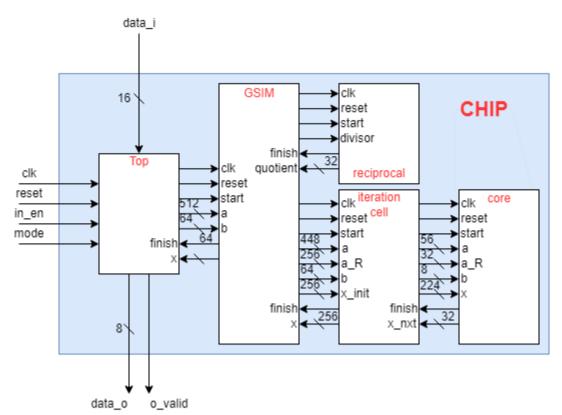
[6] 電路詳圖



除法器架構圖



運算核心架構圖



晶片架構圖

[7] 模擬結果

1. Pre-Synthesis Simulation

我們的模擬方法為使用 python 隨機產生的測試資料輸入進 CHIP 模組中,並將輸出結果另外存檔,再回到 python 進行錯誤量的運算,而錯誤量的運算方式為上述提過的,將輸入的 A 與輸出的 X(或 A^{-1})相乘後,減去輸入的 B(或單位矩陣),並將每一項的值平方後相加得到總平方誤差,在我們的晶片中此總平方誤差的平均值為 10^{-7} 。

2. Post-Synthesis Simulation

結果與 Pre-Synthesis simulation 相同,可以得知合成完的 gate level 是正確的。

3. Post-Layout Simulation

結果與 Post-Synthesis simulation 相同,可以得知 APR 完的晶片是正確的。

[8] 量測流程

- 1. 利用 Python 生成測試資料。
- 2. 儀器設置:電源供應器調整為 3.3V 直流電,並接上晶片的電源腳位
- 3. 儀器設置:訊號產生器接至晶片的輸入腳位,並產生 100MHz 的方波,輸入 clk 腳位
- 4. 儀器設置:將邏輯分析儀接至晶片的輸出腳位
- 5. 使用訊號產生器向晶片輸入所需的測試資料
- 6. 將晶片輸出經由 python 運算總平方誤差,確認結果是否與模擬結果一致

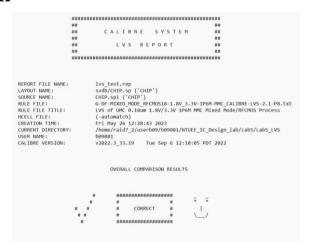
[9] 佈局驗證結果錯誤說明

1. DRC 驗證結果

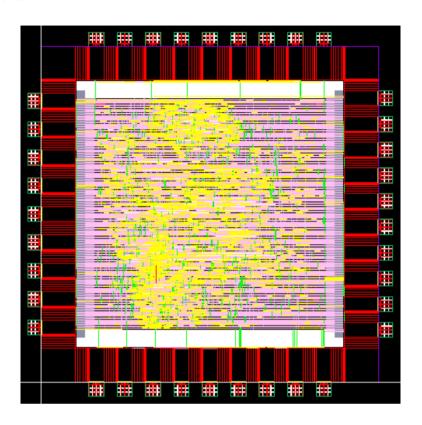
共找到錯誤如下,且皆為允許之 DRC 假錯。

```
RULECHECK RECOMMEND_4.14L ...... TOTAL Result Count = 30(3222)
RULECHECK 4.29NOTICE ...... TOTAL Result Count = 1 (1)
RULECHECK 4.14Z.NO IND PO1 ......TOTAL Result Count = 1 (1)
RULECHECK 4.20F.NO IND M1 ...... TOTAL Result Count = 1 (1)
RULECHECK 4.20G ...... TOTAL Result Count = 2 (2)
RULECHECK 4.22F.NO IND M2 ...... TOTAL Result Count = 1 (1)
RULECHECK 4.22G ...... TOTAL Result Count = 1 (1)
RULECHECK 4.24F.NO_IND_M3 ...... TOTAL Result Count = 1 (1)
RULECHECK 4.24G ...... TOTAL Result Count = 1 (1)
RULECHECK 4.26G ...... TOTAL Result Count = 1 (1)
RULECHECK 4.28G ...... TOTAL Result Count = 1 (1)
RULECHECK 4.31F ...... TOTAL Result Count = 1 (1)
RULECHECK sanity 1 ...... TOTAL Result Count = 16 (464)
RULECHECK IO5.1.W2 ...... TOTAL Result Count = 2(58)
RULECHECK IO5.1.R1 ...... TOTAL Result Count = 32 (928)
RULECHECK IO5.2.2.L1.a ... TOTAL Result Count = 16 (464)
RULECHECK IO5,2,2,L1,c ... TOTAL Result Count = 16 (464)
RULECHECK Latch.4.4.pick ...... TOTAL Result Count = 5
RULECHECK Latch.4.5.pick ...... TOTAL Result Count = 26 (641)
RULECHECK Latch.4.6.guard ...... TOTAL Result Count = 20 (570)
RULECHECK Latch.4.7 ...... TOTAL Result Count = 53 (599)
RULECHECK Latch.4.7.guard ...... TOTAL Result Count = 5 (69)
RULECHECK Latch.4.8 Latch.4.9 Latch.5.2 ... TOTAL Result Count = 1000 (53009)
RULECHECK Latch.4.10 ...... TOTAL Result Count = 8
RULECHECK Latch.5.1 ...... TOTAL Result Count = 4
RULECHECK Latch.5.5 ...... TOTAL Result Count = 98 (715)
RULECHECK 5.1A.stagger ..... TOTAL Result Count = 1 (37)
RULECHECK 5.1A1.stagger .... TOTAL Result Count = 1 (37)
```

2. LVS 驗證結果: OK



[10] 佈局平面圖

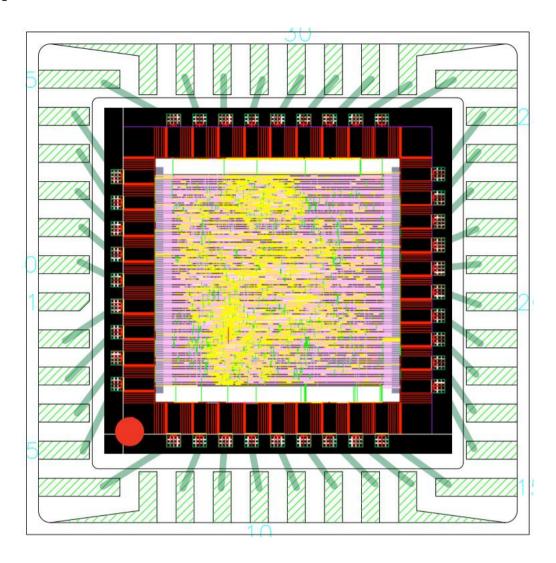


Chip Size: 1500 um x 1497 um

Power Dissipation: 9.31 mW

Max Frequency: 100MHz

[11] 打線圖



[12] 預計規格列表

Description		
Process	UMC 0.18um Mixed-Mode and RFCMOS 1.8V/3.3V 1P6M Metal Metal Capacitor Process	
Power Supply	1.8V	

Specification	Spec.	Pre-sim	Post-sim
Frequency	100 MHz	100 MHz	100 MHz
Chip size	< 1.5 mm*1.5 mm	534384 μm ²	1.500 mm*1.497 mm
Power	-	9.31 mW	9.31 mW
PADs	37	37	37

[13] 參考文獻

https://digitalsystemdesign.in/newton-raphson-reciprocal/https://en.wikipedia.org/wiki/Gauss%E2%80%93Seidel_method