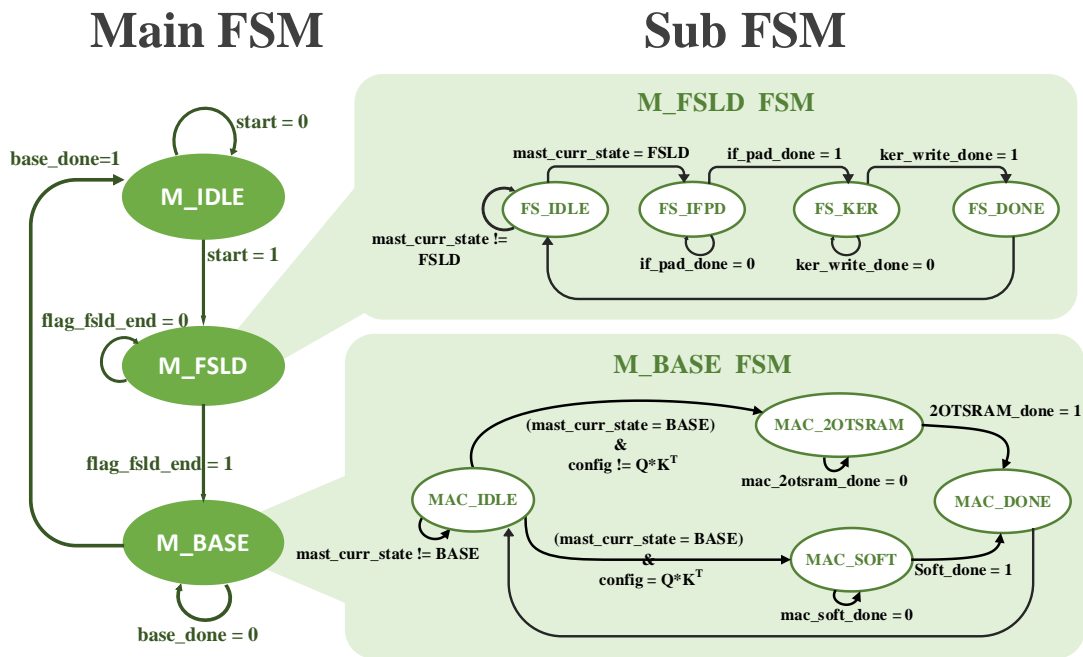




# 多頭自注意力機制與 Softmax 加速器之 硬體加速器實現

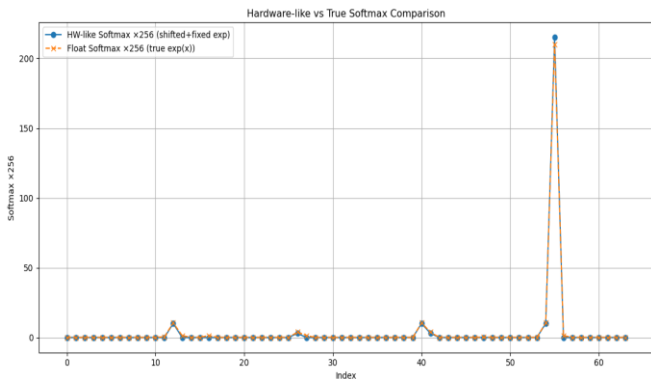
組員：電機四 蘇筠恬 黃婉芸

## FSM 有限狀態機



## Softmax 軟硬體模擬結果

圖中紅為硬體、藍為軟體的模擬結果比較。  
最大誤差約為  $5/255 \approx 1.96\%$



## 合成結果

在 TSMC 40 nm 製程下，本晶片核心面積為  $2.82 \text{ mm}^2$ 、時脈 500 MHz、總功耗 0.16 W，運算效能達 44.8 GOPS，能效 280 GOPS/W，面積效能  $15.89 \text{ GOPS/mm}^2$ 。

### TSMC40nm

Clock rate (MHz)	500 MHz
Power (W)	0.16 W
Total cell area(mm2)	2.82 mm2
Sram (KB)	74 KB
Throughput (GOPS)	44.8 GOPS
Energy Efficiency (GOPS/W)	280 GOPS/W
Area Efficiency (GOPS/ mm2)	15.89 GOPS/ m

### TSMC90nm

Clock rate (MHz)	400 MHz
Power (W)	0.36 W
Total cell area(mm2)	9.081 mm2
Sram (KB)	74 KB
Throughput (GOPS)	35.84 GOPS
Energy Efficiency (GOPS/W)	99.55 GOPS/W
Area Efficiency (GOPS/ mm2)	3.946 GOPS/ mn