## CMOS Matrix-Vector Multiplier

## ECE 533 - Final Design Review

Team Members (Group 1):

- Sk Hasibul Alam
- Tanjina Sabrin
- Md Farhan


## Outline Review

$$
\left[\begin{array}{ll}
a & b \\
c & d
\end{array}\right] \cdot\left[\begin{array}{l}
x \\
y
\end{array}\right]=\left[\begin{array}{l}
a x+b y \\
c x+d y
\end{array}\right]
$$

Input:

- $a, b, c, d, x, y \rightarrow 4$-bit unsigned each Product:
- ax, by, cx, dy $\rightarrow$ 8-bit unsigned each Output:
- $(a x+b y),(c x+d y) \rightarrow$ 9-bit unsigned

What we need:

- $4 \times 4$ Multiplier
- 8-bit Adder
- Bank of D-Latch

Adder Type:

- Carry-lookahead Adder (CLA)


## Where we left: 4-bit CLA unit

Layout



Schematic


## 4×4 Multiplier



## Things to consider:

- We need
- 5 units of CLA4
- 16 units of AND2.
- $C_{\text {out }}$ of last CLA4 is unnecessary.
- Remove OAI54321 (30 T) from that.
- P7 will be slowest (obviously).
continued...
Layout


## Schematic


continued...
4x4 Multiplier Transient Analysis (cropped)


| Transistor Count |  |  |  |
| :---: | :---: | :---: | :---: |
| Module | \#Unit | FET/unit | Total FET |
| CLA4 | 5 | 134 | 670 |
| AND2 | 16 | 6 | 96 |
| OAI54321 | -1 | 30 | -30 |
| Aggregate Count |  |  | $\mathbf{7 3 6}$ |

* We don't need $\mathrm{C}_{\text {out }}$ of last CLA4.
continued...



## Note:

$>$ Inputs are changing at 25 MHz here.

- could be increased up to 50 MHz
$>t_{p d}$ is $8 \mathrm{~ns} \rightarrow$ will affect next stage


## Dynamic D-Latch

Transmission Gate based

$\checkmark$ Slightly faster:

- Both FETs work in parallel

Clocked CMOS ( C² $^{2}$ MOS)

$\checkmark$ Slightly smaller:

- Eliminates 2 contacts

Problem with dynamic output:

- May lose value due to subthreshold leakage
- May lose value during burn-in test
continued...

Schematic


## Layout



C²$^{2}$ MOS Latch Transient Analysis (extracted)


As expected,

- Q follows D while clk is true.
- $Q$ holds $Q_{\text {old }}$ while clk is false.

But,

- $t_{p d q}$ is not bad, but not good either.


## Possible Methods of Final Design

$\left[\begin{array}{ll}a & b \\ c & d\end{array}\right] \cdot\left[\begin{array}{l}x \\ y\end{array}\right]=\left[\begin{array}{l}a x+b y \\ c x+d y\end{array}\right]$

## Assume:

Inputs are driven from D-FF. Outputs are sampled by D-FF.

Method-1

$\checkmark$ Fastest method $\left(T_{c l k}\right)$

- Both outputs concurrently available after $t_{p d}$
- Outputs ready to be sampled at next posedge of D-FF
No internal clock needed
* Enormous area required

Method-2


* Moderately fast method ( $2 \times T_{c l k}$ )
- Inputs are serially driven.
- $2^{\text {nd }}$ output $(c x+d y)$ is sampled one $T_{c l k}$ after $1^{\text {st }}$ output is sampled.
$\checkmark$ No internal clock needed
* Huge area still required
- even after $50 \%$ reduction

|  | Posedge |  |  |
| :---: | :---: | :---: | :---: |
|  | 0 | 1 | 2 |
| $\ln 1$ | $a$ | $c$ |  |
| $\ln 2$ | $x$ | $x$ |  |
| $\ln 3$ | $b$ | $d$ |  |
| $\ln 4$ | $y$ | $y$ |  |
| Out |  | $a x+b y$ | $c x+d y$ |

## continued..

Method-3


* Slowest method ( $4 \times T_{\text {clk }}$ )
- Inputs are serially driven.
* Internal clock needed
* Reset signal needed
- Reset the D-FF bank at beginning, reset it again after $(a x+b y)$ is sampled.
$\checkmark$ Small area required

|  | Posedge |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0 | 1 | 2 | 3 | 4 |  |
| $\ln 1$ | $a$ | $b$ | $c$ | $d$ |  |  |
| In2 | $x$ | $y$ | $x$ | $y$ |  |  |
| Out |  | $a x+0$ | $a x+b y$ | $c x+0$ | $c x+d y$ |  |

## Method-4


$\checkmark$ Moderately fast method $\left(2 \times T_{c l k}\right)$

- TDM is employed for D-Latch banks.
* Internal clock needed
$\checkmark$ Reset signal not needed
$\checkmark$ Smallest area required

|  | clk Level |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  | 1 | 0 | 1 | 0 |
| $\operatorname{In} 1$ | $a$ | $b$ | $c$ | $d$ |
| $\ln 2$ | $x$ | $y$ | $x$ | $y$ |
| Out | $a x+?$ | $a x+b y$ | $c x+b y$ | $c x+d y$ |

Full Design: Matrix-Vector Multiplier


## continued...

Transient Analysis


Transistor Count

| Module | \#Unit | FET/unit | Total FET |
| :---: | :---: | :---: | :---: |
| Multiplier | 1 | 736 | 736 |
| D-Latch | 16 | 6 | 96 |
| CLA4 | 2 | 134 | 268 |
| INV2 | 1 | 2 | 2 |
| Aggregate Count |  | 1102 |  |

continued...


## $\square$ Test Case:

$\left[\begin{array}{cc}14 & 5 \\ 12 & 11\end{array}\right] \cdot\left[\begin{array}{l}10 \\ 13\end{array}\right]=\left[\begin{array}{l}205 \\ 263\end{array}\right]$

- 205 = b0_1100_1101
- 263 = b1_0000_0111

Note:
$>$ clk frequency is 25 MHz here.

- can be increased up to 45 MHz
$>t_{p d}$ is 9 ns .


## Conclusion

## Problems we faced

- Only up to M3 was available.
- Too much emphasis upon equal $\mathrm{R}_{\text {rise }} \& \mathrm{R}_{\text {fall }}$
- D-FF or D-Latch?
- Placement of last CLA after D-Latch
- OAI54321: Height vs. Width


## Problems in our design

- 4x4 Multiplier
- Unnecessary $\bar{P}, \bar{G}$ from last CLA4
- Unnecessary XNOR2 gates at peripheral
- Final design
- Not running at 80 MHz as predicted


## Questions? Concerns? Comments?

