

Instituto Superior Técnico

MESTRADO INTEGRADO EM ENGENHARIA ELECTROTÉCNICA E DE COMPUTADORES

Sistemas Integrados Analógicos

Design de um Amplificador

João Bernardo Sequeira de Sá n.º 68254 Maria Margarida Dias dos Reis n.º 73099 Nuno Miguel Rodrigues Machado n.º 74236

$\acute{\mathbf{I}}\mathbf{ndice}$

1	Inti	ntrodução		1			
2	\mathbf{Ade}	denda ao <i>Middle Target</i>		1			
	2.1	1 Detecção dos erros		2			
	2.2	2 Correcção do dimensionamento		3			
	2.3	3 Demonstração de resultados		4			
		2.3.1 Resultados do dimensionamento inicia	վ	5			
		2.3.2 Resultados do dimensionamento corri	gido	5			
3	\mathbf{Pro}	rojecção do <i>Layout</i>		6			
	3.1	1 Multiplicidade e Fingers		6			
	3.2	2 Disposição dos transístores		7			
	3.3	3 Ligações internas dos blocos	Ligações internas dos blocos				
		3.3.1 Estrutura common centroid $\#1$		11			
		3.3.2 Estrutura common centroid $\#2$		11			
		3.3.3 Estrutura common centroid $\#3$		11			
		3.3.4 Estrutura common centroid $\#4$		11			
	3.4	4 Ligações externas entre os blocos		11			
4	Cor	Conclusões		12			

1 Introdução

Pretende-se projectar um amplificador folded cascode CMOS OTA de dois andares de acordo com as especificações da seguinte tabela.

Especificação	Símbolo	Valor
Tensão de Alimentação	Vdd	3.3 V
Ganho para Sinais de Baixa Amplitude	Av	70 dB
Largura de Banda	Bw	60 kHz
Margem de Fase	PM	60°
Capacidade da Carga	CL	0.25 pF
Slew-Rate	SR	200 V/μs
Budget da Corrente	IDD	400 μΑ
Área de <i>Die</i>	/	0.02 mm^2

Tabela 1: Características do amplificador a projectar.

O circuito de ponto de partida para a realização do projecto é apresentado de seguida.



Figura 1: Circuito do amplificador a projectar.

${\bf 2} \quad {\bf Adenda\ ao}\ {\it Middle\ Target}$

Esta secção foi acrescentada ao relatório final no intuito de corrigir os resultados obtidos e apresentados no relatório anterior, o do *middle target*. Como referenciado, pretende-se projectar um amplificador folded cascode CMOS OTA de dois andares de acordo com as especificações da Tabela 1.

2.1 Detecção dos erros

Foram identificados vários erros no relatório intermédio que comprometem os resultados apresentados anteriormente. A primeira correcção foi referente ao *schematic* do *testbench* que permite simular o circuito em testes de resposta AC. Foi colocado um *switch* que simula a bobine - provoca um circuito aberto para um regime AC e curto-circuito para um regime DC. Foi também alterada a amplitude do sinal de entrada de 3.3 V para 1.6 V, sendo que esta alteração garante que os transístores não saem da saturação. De seguida pode-se comparar o novo *testbench* com o anterior.

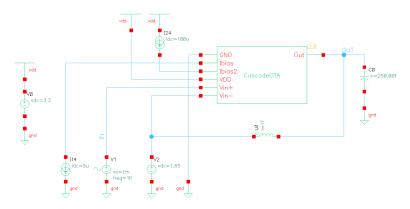


Figura 2: Schematic do testbench anterior que permite simular o circuito em testes de resposta AC.

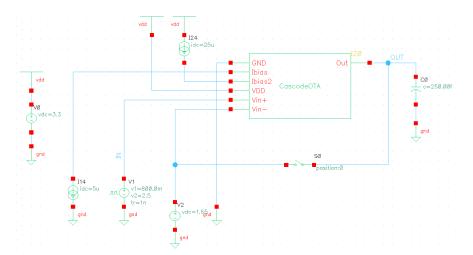


Figura 3: Schematic do novo testbench que permite simular o circuito em testes da slew-rate e de resposta transiente, DC e AC.

Outro erro identificado é referente ao cálculo da *slew-rate*. No relatório intermédio, o resultado da *slew-rate* era relativo só ao flanco de descida, sendo necessário demonstrar para os dois flancos - subida (equação 2.1) e descida (equação 1.2).

$$slewRate(VT("/OUT") 1 nil 2 nil 10 90 nil "time")$$
(2.1)

A equação slewRate é obtida com recurso à calculadora do Cadence, escolhendo o sinal pretendido, que, neste caso, é a saída do *cascode* OTA, VT("/OUT"). De seguida escolhe-se a posição inicial e final

para o cálculo da *slew-rate* - foi definido que para o flanco de descida começa-se a calcular desde 2 V até 1 V e no flanco de subida o cálculo é o inverso, de 1 V até 2 V. O intervalo anteriormente referido foi escolhido de forma a calcular a *slew-rate* na zona onde os transístores se encontram na saturação.

2.2 Correcção do dimensionamento

Com os erros anteriormente referidos corrigidos, o circuito apresentado na entrega intermédia falhava em algumas especificações: ganho para sinais de baixa amplitude, margem de fase e largura de banda. Na secção seguinte apresenta-se os resultados das simulações de Monte Carlo e *corners*.

Com o intuito de corrigir as dimensões, partiu-se de um critério inicial - obter transístores de dimensões mais reduzidas e manter o rácico W/L obtido no relatório anterior. Em primeiro lugar, verificou-se que com a nova expressão do cálculo da slew-rate obtém-se valores mais elevados, o que significa que se pode reduzir as dimensões dos transístores M_9 e M_{10} .

De seguida pretendeu-se melhorar a margem de fase e, analisando o resultado das especificações pretendidas, verificou-se que era necessário aumentar a largura de banda de forma a obter uma margem de fase melhor sem alterar significativamente os resultados das outras especificações. Começou-se por analisar as dimensões das capacidades do circuito e verificou-se que a capacidade referente aos transístores M_7 e M_8 é a capacidade que mais influencia a margem de fase. Mas, como já foi referido no relatório anterior, os transístores M_7 e M_8 também afectam o ganho e a largura de banda, como se pode ver pelas expressões seguintes.

$$A_v = g_{m_1} R_o = g_{m_1} \left[\left(g_{m_4} r_{o_4} \left(r_{o_2} / / r_{o_{10}} \right) \right) / \left(g_{m_6} r_{o_6} r_{o_8} \right) \right]; \tag{2.3}$$

$$f_p = \frac{1}{2\pi C_L R_o} = \frac{1}{2\pi C_L \left[\left(g_{m_4} r_{o_4} \left(r_{o_2} / / r_{o_{10}} \right) \right) / \left(g_{m_6} r_{o_6} r_{o_8} \right) \right]}; \tag{2.4}$$

Assim sendo, e analisando as expressões decidiu-se inicialmente diminuir as dimensões dos transístores M_7 e M_8 , o que fez aumentar a largura de banda como também a margem de fase. Com o intuito de melhorar as especificações obtidas diminui-se as dimensões dos transístores M_3 e M_4 até atingir uma dimensão de L miníma de $0.85~\mu m$. Como a dimensão de L dos transístores M_5 e M_6 é menor que $0.85~\mu m$, aumentou-se as dimensões com o intuito de manter o critério de um L mínimo de $0.85~\mu m$.

Com estas alterações consegui-se obter todas a especificações excepto para o ganho de sinais de baixa amplitude. Analisando a expressão anterior do ganho, verifica-se que para aumentar o ganho sem alterar as restantes especificações, seria necessário aumentar o g_{m_1} . Como o valor de g_m é directamente proporcional a I_D , com aumento de W existe um aumento de I_D e de g_m . A expressão seguinte mostra como W influencia I_D :

$$g_m \propto I_D = k_P \times \left(\frac{W}{L}\right) \times V_{OD}^2 \times (1 + \lambda V_{DS}).$$
 (2.5)

Para aumentar o g_{m_1} é necessário alterar o rácio W/L, ou seja, o rácio de M₁ e M₂ sofreu um aumento de 132%, passou de 43 para 100.

Assim, o circuito final relativamente à entrega intermédia é:

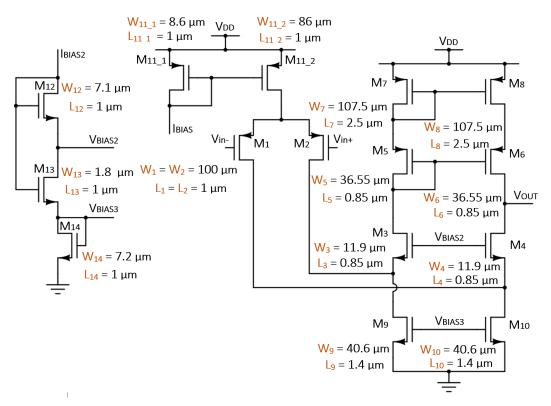


Figura 4: Circuito final da entrega intermédia.

Tabela 2: Dimensões dos transístores que constituem o amplificador.

Transístores	W [μm]	L [μm]	Rácio W/L
M1 e M2	100	1	100
M3 e M4	11.9	0.85	14
M5 e M6	36.55	0.85	43
M7 e M8	107.5	2.5	43
M9 e M10	40.6	1.4	29
M11_1	8.6	1	8.6
M11_2	8.6	1	8.6
M12	7.1	1	7.1
M13	1.8	1	1.8
M14	7.2	1	7.2

Com este circuito as especificações finais obtidas para o circuito com recurso à simulação da Figura 3 são apresentadas na seguinte tabela.

simular e retirar valores e graficos

2.3 Demonstração de resultados

Nesta secção apresenta-se os resultados do dimensionamento do relatório intermédio e do novo dimensionamento anteriormente referido, para simulações de Monte Carlo e de *corners*.

2.3.1 Resultados do dimensionamento inicial

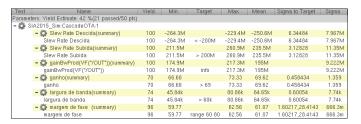


Figura 5: Simulação de Monte Carlo para 50 pontos.



Figura 6: Resultados obtidos para as 3 primeiras simulações de Monte Carlo.

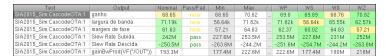


Figura 7: Simulação por corners.

2.3.2 Resultados do dimensionamento corrigido

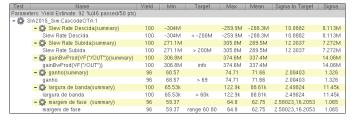


Figura 8: Simulação de Monte Carlo para 50 pontos.

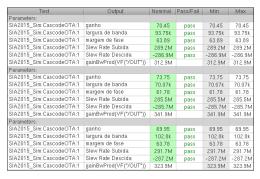


Figura 9: Resultados obtidos para as 3 primeiras simulações de Monte Carlo.

Test	Output	Nominal	Pass/Fail	Min	Max	corner(WP	WS	WO	WZ
SIA2015_Sim:CascodeOTA:1	ganho	70.97	pass	70.06	73.82	disab	72.16	71.89	70.06	73.82
SIA2015_Sim:CascodeOTA:1	largura de banda	96.3k	pass	76.45k	97.18k	disab	96.77k	76.45k	97.18k	77.62k
SIA2015_Sim:CascodeOTA:1	margem de fase	63.89		59.9	65.11	disab	64.16	61.58	65.11	59.9
SIA2015_Sim:CascodeOTA:1	Slew Rate Subida	292.4M	pass	277.3M	303.8M	disab	298.5M	284.9M	277.3M	303.8M
SIA2015_Sim:CascodeOTA:1	Slew Rate Descida	-291.5M	pass	-301.5M	-283.4M	disab	-287.3M	-288.4M	-283.4M	-301.5M
SIA2015_Sim:CascodeOTA:1	gainBwProd(VF("/OUT"))	341.4M		301.2M	393.3M	disab	393.3M	301.2M	310.1M	382.1M

Figura 10: Simulação por corners.

comentar diferenças obtidas

3 Projecção do Layout

3.1 Multiplicidade e Fingers

Para projectar o layout do circuito da Figura 4 optou-se por dividir os transístores de grandes dimensões com recurso a duas técnicas - multiplicidade e fingers. A técnica de fingers corresponde a um arranjo específico do transístor com n gate fingers em que as difusões da source e do drain são partilhadas. Se se tiver n fingers haverá então n+1 difusões. A multiplicidade é quando se faz uma ligação em paralelo de múltiplos dispositivos MOS, sendo que o agregado deles corresponde a um só transístor.

Do ponto de vista da definição correspondem ao mesmo, mas são de facto duas maneiras diferentes de pensar na paralelização de transístores. Com o recurso a *fingers* tem-se uma única célula com o transístor completo com todos os *fingers*, útil para quando se quer uma célula mais compacta, enquanto na multiplicidade tem-se tantos transístores quanto a multiplicidade indicar. De facto, é possível conjugar as duas técnicas, ou seja, cada dispositivo MOS da multiplicidade pode ser feito com vários *fingers*.

Para o trabalho em causa optou-se por usar as duas técnicas teoricamente idênticas para adquirir mais experiência e para verificar as diferenças que têm ao nível de implementação prática no Cadence.

Estas práticas são aplicadas também de forma a evitar efeitos nocivos de *mismatches*, ou seja da vulnerabilidade aos gradientes de parâmetro. Ao minimizar a área efectiva dos circuitos protege-se assim o dispositivo destes efeitos.

Sendo assim deu-se especial atenção aos pares de transístores em que se considerou que existe uma maior sensibilidade a *mismatches*. São este os pares M_1/M_2 , M_7/M_8 , M_9/M_{10} e M_{11_1}/M_{11_2} .

Começou-se em primeira instância por aplicar multiplicidade, tentando aproximar ao máximo os tamanhos dos pares que estão ligados entre si, seguido de uma redução geral dos transístores de todo o circuito.

Em adição a multiplicidade optou-se por usar também a técnica fingers no par M_7/M_8 . Já no par $5/M_6$ aplicou-se apenas a técnica fingers ficando-se assim com tamanhos da mesma ordem em ambos os pares para além de uma forma compacta facilitando as tarefas de escolher uma disposição para os transístores e efectuar as ligações internas.

Como o rácio de espelhamento é muito sensível a diferenças nos parâmetros do espelho de corrente, ou seja a mismatches, nos pares de transistores M_{11_1}/M_{11_2} . Dado isto foi aplicada multiplicidade no

melhorar explicação de fingers e multiplicidade a seguir transístor M_{11_2} de forma a que tivesse o mesmo tamanho que o M_{11_1} sendo assim o par afectado de igual forma pelos gradientes de parâmetros.

Para o par M_9/M_{10} , como são transístores NMOS e se optou por colocar num grupo apenas de NMOS, como será explicado à frente, houve apenas necessidade de reduzir os seus tamanhos através de multiplicidade para valores menos sensíveis a *mismatches*.

Para os restantes cuja sensibilidade a variação de parâmetros não é tão importante quanto nos pares supramencionados usou-se multiplicidade de forma a obter-se as dimensões desejadas e também que facilitassem as estruturas do layout.

Assim, na tabela seguinte encontra-se uma descrição de como são constituídos os vários transístores do circuito.

Transístores	Multiplicidade	Número de gates	Wtotal [μm]	Wsingle gate [μm]	L [μm]	Descrição
M1 e M2	8	1	100	12.5	1	8 transístores, cada um com 1 <i>gate</i>
M3 e M4	2	1	11.9	5.95	0.85	2 transístores, cada um com 1 gate
M5 e M6	Ms e M6 1 4		36.6	9.15	0.85	1 transístor com 4 <i>gates</i>
M7 e M8	2	4	107.5	13.45	2.5	2 transístores, cada um com 4 gates
M9 e M10	4	1	40.6	10.15	1.4	4 transístores, cada um com 1 gate
M11_1	1	1	8.6	8.6	1	1 transístor com 1 <i>gate</i>
M11_2	10	1	8.6	0.86	1	10 transístores, cada um com 1 gate
M12	1	1	7.1	7.1	1	1 transístor com 1 <i>gate</i>
M13	1	1	1.8	1.8	1	1 transístor com 1 <i>gate</i>
M14 1		1	7.2	7.2	1	1 transístor com 1 <i>gate</i>

Tabela 3: Dimensões e características dos transístores do amplificador.

3.2 Disposição dos transístores

Depois de se definir como estão estruturados os vários transístores é importante definir como se encontram dispostos no layout. A topologia básica que foi utilizada é a de common centroid. Através desta técnica consegue-se garantir um melhor matching entre dois transístores iguais e em que se pretende um comportamento semelhante. De facto, o common centroid é utilizado para se garantir que, e.g., um amplificador diferencial tenha um sinal de modo comum próximo de 0 e, como tal, um CMRR maior.

Para o circuito em causa foram definidos 4 blocos principais sobre os quais se definiu uma estrutura common centroid:

- espelho de corrente básico que é polarizado em corrente com I_{BIAS} (equivalente ao Bloco 1 da Figura 1) estrutura common centroid #1;
- transístores do par diferencial (Bloco 2 da Figura 1) estrutura common centroid #2;
- espelho de corrente cascode básico do tipo PMOS (Bloco 3 da Figura 1) estrutura common centroid #3;
- transístores do tipo NMOS estrutura common centroid #4;.

A escolha destes blocos foi feita com base no mismatch. È essencial que entre os transístores do par diferencial não haja diferenças de comportamento, pois se houver o circuito fica desequilibrado. Os dois

espelhos de corrente existentes também não devem ter *mismatch*, pois pretende-se um espelhamento correcto. Como os transistores NMOS eram os mais pequenos e como não necessitam de estar dentro de um poço optou-se por juntar todos em apenas um bloco de forma a facilitar também a obtenção de *common centroid* como será falado à frente.

De seguida apresenta-se as várias estruturas common centroid definidas:

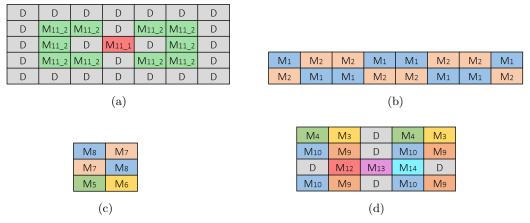


Figura 11: Estrutura common centroid #1 (a), #2 (b), #3 (c) e #4 (d).

Nas figuras anteriores verifica-se a existências de transístores representados pela letra D - transístores dummies. Este tipo de transístores serve para se garantir que os restantes transístores têm a mesma vizinhança, ou seja, que à sua volta vêem o mesmo. Veja-se o seguinte exemplo em que se consideram dois transístores - 1 e 2. O transístor 1 está definido com uma multiplicidade de 1 e o transístor 2 com uma multiplicidade de 8. Na figura seguinte encontra-se um exemplo de estrutura common centroid para esse caso.

2	2	2		
2	1	2		
2	2	2		

Figura 12: Exemplo de estrutura common centroid com dois transístores.

O transístor 1 tem como vizinhança 4 transístores - para cima, para baixo, para a esquerda e para a direita. No entanto, qualquer transístor 2 tem uma "falha" naquilo que vê, pois não tem "vizinhos" nalgumas direcções. Assim, para se resolver este problema colocam-se transístores dummies nos sítios onde há "falhas".

	D	D	D	
D	2	2	2	D
D	2	1	2	D
D	2	2	2	D
	D	D	D	

Figura 13: Exemplo de estrutura common centroid com dois transístores e também transístores dummies.

De facto, os transístores *dummies* não necessitam de ter as mesmas dimensões que os restantes e o circuito da figura anterior pode ser definido sa seguinte maneira.

	D		D	
D	2	2	2	D
D	2	1	2	D
D	2	2	2	D
	D	D	D	

Figura 14: Exemplo de estrutura common centroid com dois transístores e também transístores dummies de dimensões menores.

Relembrando as estruturas da Figura 11, verifica-se que a estruturas #1 e #4 têm dummies.

Na caso da estrutura #1, tinha-se à partida 11 transístores de dimensões iguais após a aplicação de multiplicidade. Como se quer garantir que todos os transístores observam as mesmas vizinhanças minimizando o efeito de mismatches, fez-se uso de dummies para ocupar as lacunas que iriam de outra forma ficar vazias. O ponto de partida foi assim o transístor M_{11_1} em que não se aplicou multiplicidade rodeando-o dos 10 transístores que constituem M_{11_2} . Após isto observou-se que seriam assim necessário 24 dummies. Obtém-se assim uma estrutura rectangular de 5X7 em que todos os M_{11_1} como M_{11_2} observam as mesmas vizinhanças.

Para a estrutura #4 o procedimento foi similar. Tinha-se á partida os 3 transístores de polarização M_{12} , M_{13} e M_{14} aos quais não foi aplicada nem multiplicidade nem fingers pois já têem tamanhos bastante reduzidos. Para que se pudesse garantir que eles tinham as mesmas vizinhanças teriam que ser rodeados de dummies. Como no entanto se tratam de transístores do tipo N e tinha sido decido colocar todos os transístores deste tipo no mesmo bloco bastou adicionar-se os pares M_3/M_4 e M_9/M_{10} , colocando-os de forma alternada obtendo-se o presente na Figura 11 (d). Como todos os transístores presentes neste bloco têm tamanhos muito díspares fica-se uma estrutura final que apresenta algumas áreas vazias.

falar sobre como os buracos do bloco

3.3 Ligações internas dos blocos

Uma vez definidos como estão estruturados os transístores e como estão dispostos efecutam-se as ligações internas dos 4 blocos da Figura 11.

Existem várias regras para construir um *layout*, sendo que uma delas impõe um espaço mínimo para separar as diferentes máscaras. Para evitar isto, recorre-se a uma opção existente no Cadence que notifica o utilizador quando este não cumpre as margens mínimas.

As ligações entre gates de transístores que estejam próximos são feitas a partir da camada condutora poly. Porém, se os transístores estiverem afastados é preferível efectuar a ligação com recurso a Metal 1 e a contactos do tipo P1_C, que efectuam a ligação entre poly e Metal 1. Esta solução é preferível para esses casos pois a poly tem uma resistividade elevada e colocar demasiado dessa camada faz aumentar a resistividade geral do circuito. De facto, na tabela seguinte apresenta-se a resistividade das diversas máscaras existentes no fabrico e verifica-se que a poly (poly 1) tem uma resistividade bem superior face a camadas como Metal 1 e Metal 2. Verifica-se também que a resistividade de um contacto do tipo P1_C é elevada, mas que compensa, uma vez que para efectuar uma ligação correcta entre duas gates bastaria colocar no máximo 2 contactos.

Tabela 4: Resistividade das diversas máscaras de fabrico (a) e resistividade dos diversos tipos de contactos (b).

Sheet resistance Layer Ω/\Box metal4 $\overline{R_{sm4}}$ 0.05 R_{sm3} metal3 0.05 R_{sm2} 0.08 metal2 metal1 R_{sm1} 0.08poly1 R_{sp} R_{sp2} poly2 50

 R_{sdn}

 R_{sdp}

80

150

 n^+ diff.

 p^+ diff.

(a)

Contact resistance	
Layer-layer	Ω/cnt
metal4-metal3	R_{via3} 3
metal3-metal2	R_{via2} 1.5
metal2-metal1	R_{via} 1.5
metal1-poly1	$R_{cp} = 5$
metal1-n ⁺ diff.	R_{cdn} 40
metal1-p ⁺ diff.	R_{cdp} 90

(b)

Tomou-se a decisão de, sempre que possível, efectuar as ligações com Metal 1 na horizontal e com Metal 2 na vertical. Procura-se aplicar esta *rule of thumb* sempre que possível, sendo que apenas não é cumprida quando se verifica que acaba por tornar mais difícil o *design* do *layout* ou que conduz a uma área maior porque é necessário ajustar o espaçamento entre transístores para acomodar uma ligação em Metal 1 ou Metal 2 consoante o caso.

Para ligar Metal 1 a Metal 2 é usada uma via. De referir que, optou-se também, sempre que possível, por colocar vias e contactos duplos, para que houvesse um de backup. À semelhança do que se referiu anteriormente, isto é feito quando não dificulta o design do layout.

3.3.1 Estrutura common centroid #1

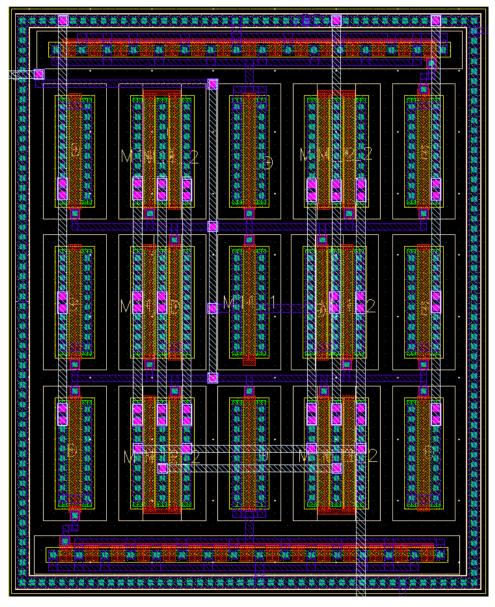


Figura 15: Layout da estrutura que corresponde ao espelho de corrente básico que é polarizado em corrente com I_{BIAS} .

- 3.3.2 Estrutura common centroid #2
- 3.3.3 Estrutura common centroid #3
- 3.3.4 Estrutura common centroid #4
- 3.4 Ligações externas entre os blocos

4 Conclusões