

6T SRAM电路结构与关键技术 节点中的工艺流程简述

本书旨在让光刻初学者在短时间内了解光刻的基本知识,除了前两章中的工厂知识和光刻的基础知识,还需要了解晶体管的基本结构、简单的工艺流程以及光刻工艺处于工艺流程中的具体位置。了解工艺流程,一方面可以更好地完成部门间的协同工作;另一方面,在出现问题时可以更快地找到问题的原因与责任归属。

光刻工艺是整个工艺流程中最重要的一环。如图 3.1 所示,本章先介绍光刻工艺处于整个工艺流程中的位置,然后介绍静态随机存取存储器(Static Random Access Memory, SRAM)的电路结构和基本工作原理,最后介绍 SRAM 中三种主要的晶体管种类:平面晶体管(Planar Transistor)、鳍型场效应晶体管(Fin Field Effect Transistor, FinFET)和互补场效应晶体管(Complementary Field Effect Transistor, CFET)的基本结构与几个关键技术节点下 SRAM 的简单工艺流程。

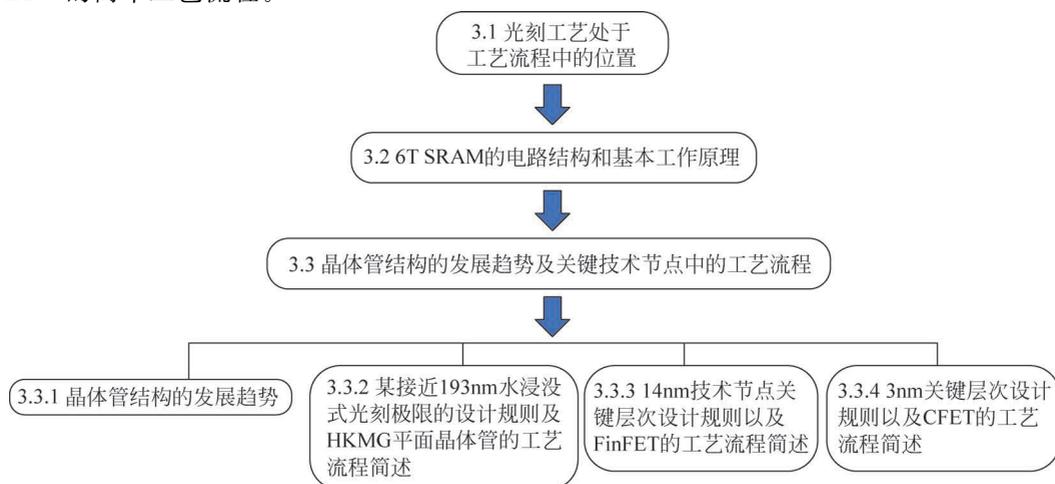


图 3.1 本章主要内容简介

3.1 光刻工艺处于工艺流程中的位置

图 3.2 展示了集成电路中光刻工艺所处位置以及简单的工艺流程。

(1) 首先,需要进行薄膜沉积(薄膜 1),根据需求选择薄膜的种类和厚度,包括氧化硅、氮化硅等复杂的无机物薄膜。

(2) 接下来,需要利用轨道机与光刻机完成涂胶、曝光、显影等一系列工艺流程(具体见 4.3 节),在光刻胶上形成图形。图中只展示了光刻胶,一般还会使用 1~2 层抗反射层(见 6.7.5 节)。

(3) 通过光刻工艺定义出图形之后,光刻胶可以作为阻挡层进行下一步的离子注入或者刻蚀工艺:离子注入工艺定义掺杂区(如 P 型、N 型等),刻蚀工艺是将光刻定义的图形保真传递到下层的抗反射层以及薄膜 1 上。

(4) 随后将多余的光刻胶去除。对于离子注入工艺,去胶后需要一个退火的过程,以激活注入的离子。

(5) 对于刻蚀保真传递后的图形,需要继续沉积薄膜材料 2,例如,后段连线层的金属填充材料。

(6) 通过使用化学机械平坦化(Chemical Mechanical Planarization, CMP)工艺,去除多余的薄膜 2 材料。到了这一步,集成电路工艺中这一层次的工艺流程就基本结束了。

(7) 接下来开始下一个层次的工艺流程,还是先从无机物的薄膜沉积(膜层 3)开始,再进行后续的光刻、离子注入或者刻蚀、填充、CMP 等工艺流程。完成所有层次后,还会生长保护层(Passivation)并对芯片进行封装、测试等。

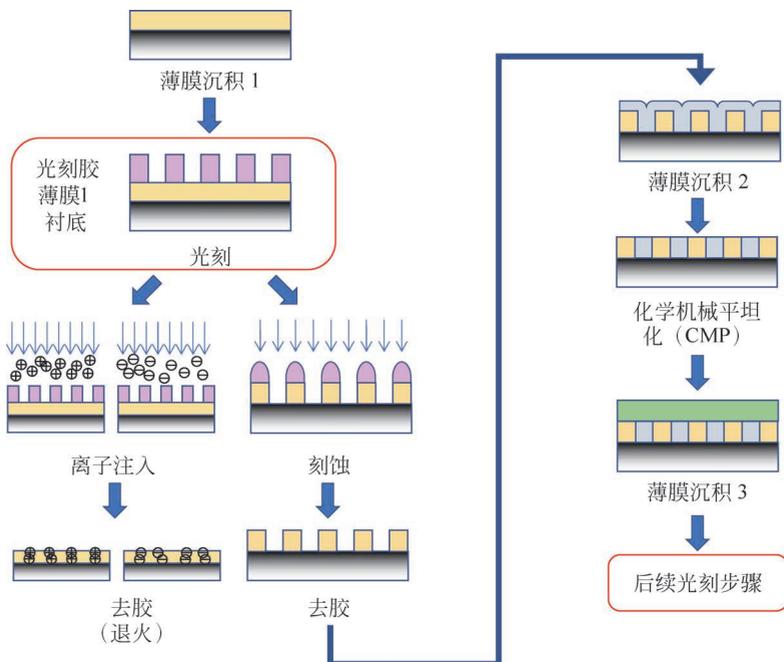


图 3.2 光刻及后续简单工艺流程示意图

可见,无论是哪个技术节点,都需要先利用光刻工艺定义出图形之后,才能进行后续的一系列工艺流程。另外,成熟的光刻技术及其他制造工艺,如刻蚀、薄膜沉积、离子注入等工艺,

低电位,具体见“读”操作。 M_5 和 M_6 称为传送门(Pass Gate,PG)晶体管,作为开关晶体管可以实现下文中位线(Bit Line,BL)对 SRAM 状态的读写操作。

(3) 一个 PU 和一个 PD 晶体管组成一个反相器,反相器就是“非门”电路,其输出是输入的逻辑非。因此 $M_1 \sim M_4$ 可以组成两个交叉耦合的反相器,用来存储这一比特的数据,具体如下。

① M_1 和 M_2 组成一个反相器, M_3 和 M_4 组成另外一个反相器。

② 可以看出,两个反相器的输入和输出交叉互连,即 M_1 和 M_2 组成的反相器的输出 \bar{Q} 作为 M_3 和 M_4 组成的反相器的输入;而 M_3 和 M_4 组成的反相器的输出 Q 作为 M_1 和 M_2 组成的反相器的输入。

③ 这种输入和输出交叉互连的方式,可以实现一比特的存储。即当一个反相器的输入为“1”时,输出即为“0”,此输出信号输入另外一个反相器中,最后输出信号“1”。

对于 SRAM 来说,只要可以提供电源(V_{DD}),信号“1”可以在反相器中循环出现,而不需要像 DRAM 一样定期刷新电路。除了晶体管外,图 3.3 中的 SRAM 电路中还包含字线(Word Line,WL),控制字节的存取;位线(Bit Line,BL 或 \bar{BL}),用来读取或者写入一个 SRAM 基本单元的状态。

无论对一个 SRAM 基本单元进行何种操作,字线都需要处于高电位状态,将作为此 SRAM 基本单元的两个控制开关 M_5 和 M_6 开通,通过这两个开关(PG),将 SRAM 中的两个反相器和位线连通。连通了位线,才能对 SRAM 进行“读”或者“写”数据的操作。如果字线没有处于高电位状态,相当于开关 M_5 和 M_6 断开,位线与两个反相器中的 4 个晶体管 $M_1 \sim M_4$ 隔断,两个反相器继续保持其原有的状态,即“空闲”(Standby)状态。此外,还可以对 SRAM 基本单元进行“读”和“写”的操作^[1-2]。接下来,分别介绍如何进行“读”和“写”。

3.2.2 对 SRAM 单元进行“读”的操作

1. 读取数据“1”

在进行读的操作时,要假设 SRAM 中已经有存储数据, Q 为逻辑“1”, \bar{Q} 为逻辑“0”。同时,需要先将两个位线 BL、 \bar{BL} 预设为等电位的,一般为高电位(逻辑“1”),再将字线连通高电位。而“读”的这一过程是为了将 Q 处保存的逻辑“1”经过预设高电位的位线 BL 输出,而将预设高电位的位线 \bar{BL} 中的逻辑“1”变成逻辑“0”。如图 3.4 所示,读取 SRAM 为“1”状态的具体原理如下。

(1) 两个位线 BL、 \bar{BL} 预设逻辑“1”之后,字线 WL 加高电位。由于 NMOS 晶体管 M_1 和 M_5 的栅极的输入都是高电平, M_1 和 M_5 导通。同时,PMOS 晶体管 M_2 的栅极的输入也是高电平,因此 M_2 无法导通。而 \bar{BL} 直接通过 M_5 和 M_1 接地(V_{SS}),从预设的逻辑“1”变成“0”,由于通过 M_1 可以拉低电位,因此 M_1 称为下拉(PD)晶体管。

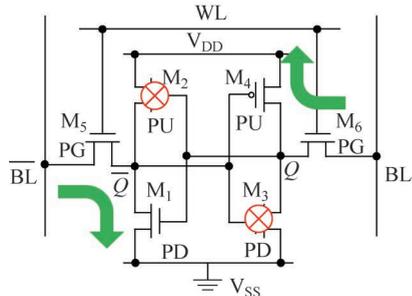


图 3.4 对 SRAM 存储数据为“1”的状态进行“读”操作的示意图

(2) 在 BL 一侧,NMOS 晶体管 M_6 导通,同时由于 PMOS 晶体管 M_4 的栅极输入的是低电位(\bar{Q} 为“0”),因此 M_4 也导通。由于 NMOS 晶体管 M_3 的输入是低电平,因此无法导通。而 BL 直接通过 M_6 和 M_4 连接到电源(V_{DD})代表的逻辑“1”,由于通过 M_4 可以拉高电位,因此 M_4 称为上拉(PU)晶体管。因此,完成了位线 BL 将 Q 处存储的“1”读出的操作。

2. 读取数据“0”

相反地,若 SRAM 中初始存储的是“0”,即 Q 为逻辑“0”,则 \bar{Q} 为逻辑“1”。如图 3.5 所示,读取 SRAM 为“0”状态的具体原理如下。

(1) 当字线 WL 加高电位,在 \bar{BL} 一侧,NMOS 晶体管 M_5 导通,同时由于 PMOS 晶体管 M_2 的栅极的输入是低电位(Q 为“0”),因此 M_2 也导通,而 \bar{BL} 直接通过 M_5 和 M_2 连接到电源(V_{DD})代表逻辑“1”,与 M_4 类似,由于通过 M_2 可以拉高电位,因此 M_2 也称为 PU 晶体管。

(2) 在 BL 一侧,NMOS 晶体管 M_3 和 M_6 导通, BL 直接通过 M_6 和 M_3 接地(V_{SS}),从预设的逻辑“1”变成“0”,从而完成位线 BL 将 Q 处存储的“0”读出的操作。与 M_1 类似,通过 M_3 可以拉低位线电位,因此 M_3 也称为 PD 晶体管。

综上,在“读”操作之前,将 BL 和 \bar{BL} 设置为相同的高电位,一旦两者之间有一个电位被拉低,实现了两个位线之间的电位差,读取信号的放大器就可以识别出哪个位线是“1”,哪个位线是“0”,完成“读”的操作。

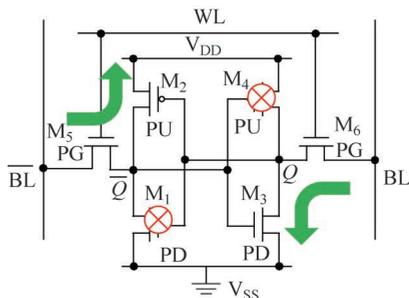


图 3.5 对 SRAM 存储数据为“0”的状态进行“读”操作的示意图

3.2.3 对 SRAM 单元进行“写”的操作

1. 写入数据“1”

位线 BL (\bar{BL})的作用是通过开关晶体管 M_5 、 M_6 将数据从一个 SRAM 基本单元中读出,或者将数据写入一个 SRAM 基本单元。前面介绍了如何读出数据,本节简单讲解写入数据的基本原理。在将数据“写”入 SRAM 之前,需要先把要写入的状态加载到位线上。即设置 BL 和 \bar{BL} 的电压差,使得两个反相器的输出 Q 和 \bar{Q} 转变状态。具体原理如下。

(1) 假设两个反相器的输出 Q 和 \bar{Q} 分别为“0”和“1”的状态。若希望将“1”写入这个 SRAM 单元,则 BL 需要设为“1”(高电位),而 \bar{BL} 设为“0”(低电位)。随后将字线连通高电位,此时晶体管导通状态如图 3.6(a)所示。

(2) 由于 \bar{BL} 为低电位, \bar{Q} 的电位会逐渐下降,这会导致 M_4 逐渐开启,而 M_3 逐渐关掉, Q 的电位会逐渐被抬升到“1”,最后显示的状态如图 3.6(b)所示,同时 M_2 关掉而 M_1 开启,数据翻转完成,即完成了写入的动作。

注意,为了保证数据顺利写入, \bar{Q} 的电位必须能够下降,所以 $PU(M_2)$ 的电流需要小于 PG 晶体管的电流(M_5),这样 PU 的电源(V_{DD})就不足以拉升 \bar{Q} 的电位,这样一来 \bar{Q} 的电位就由 \bar{BL} 的电位来决定。

2. 写入数据“0”

同样地,如果 SRAM 本来存储的数据为“1”(Q),需要将“0”写入此基本单元,即将 Q 和 \bar{Q} 的状态从“1”和“0”翻转为“0”和“1”,同样需要先把即将写入的状态加载到位线上。具体原理如下。

(1) 此时假设两个反相器的输出 Q 和 \bar{Q} 分别为“1”和“0”的状态。因为希望将“0”写入这个 SRAM 单元,那么 BL 需要设为“0”(低电位),而 \bar{BL} 设为“1”(高电位)。随后将字线连通高电位,此时晶体管导通状态如图 3.7(a)所示。

(2) 由于 BL 为低电位, Q 的电位会逐渐下降,这会导致 M_2 逐渐开启,而 M_1 逐渐关掉, \bar{Q}

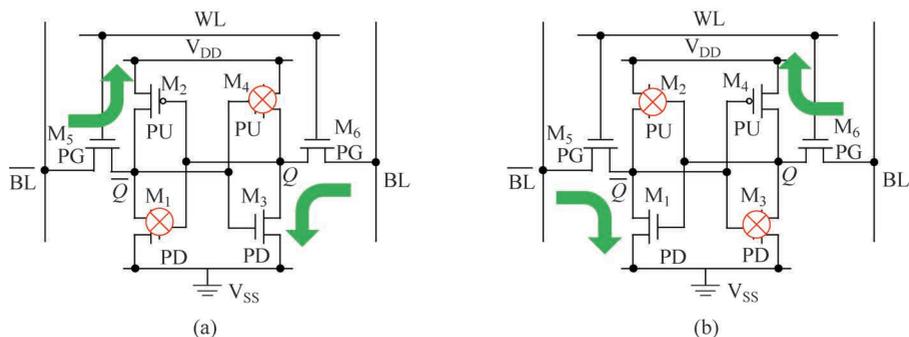


图 3.6 将数据“1”写入一个 SRAM 基本单元的示意图：(a) 初始数据状态为“0”(\bar{Q})；(b) 成功写入“1”(Q)之后的状态

的电位会逐渐被抬升到“1”，最后显示的状态如图 3.7(b)所示，同时 M_4 关掉而 M_3 开启，数据翻转完成，即完成了数据“0”写入的操作。

注意，为了保证数据顺利写入， Q 的电位必须能够下降，因此 $PU(M_4)$ 的电流需要小于 PG 晶体管的电流(M_6)，这样 PU 的电源(V_{DD})就不足以拉升 Q 的电位，这样一来 Q 的电位就由 BL 的电位来决定。写入完成之后，拉低字线电位，即可完成数据锁存。

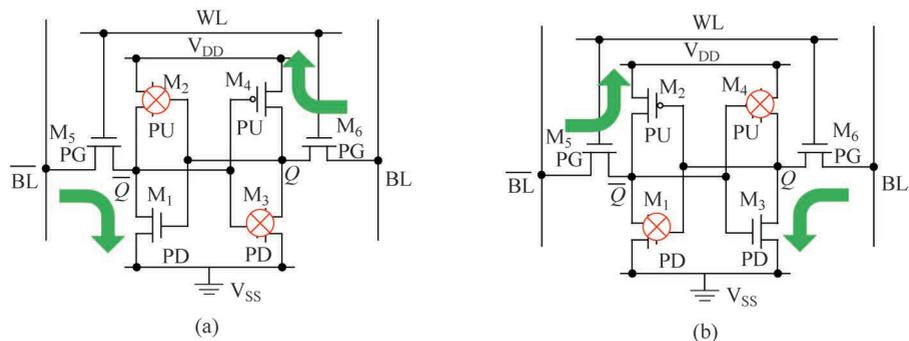


图 3.7 将数据“0”写入一个 SRAM 基本单元的示意图：(a) 初始数据状态为“1”(Q)；(b) 成功写入“0”(\bar{Q})之后的状态

综上所述，无论是“读”还是“写”的操作，字线都需要设为高电位，将作为开关的 M_5 、 M_6 与位线连接。在“读”时，需要先将 BL 和 \bar{BL} 设成等电位，然后将字线设为高电位后，利用 SRAM 内部反相器中存储的数据 Q 使得 BL 和 \bar{BL} 形成电位差（由内影响外），再利用信号放大器将此电位差放大，完成“读”的操作。在“写”时，通过外部输入的有电位差的 BL 和 \bar{BL} 改变 SRAM 中存储数据 Q 的状态（由外影响内），完成“写”的操作。

3.3 晶体管结构的发展趋势及关键技术节点中的工艺流程

3.3.1 晶体管结构的发展趋势

前面简单介绍了常见的一种 6T SRAM 晶体管的基本电路结构和工作原理，本节会介绍如何在硅片上实现这样的晶体管结构。当然，实际工艺研发之前，需要先将电路结构转换为版图设计，具体描述可见 4.6 节。随着技术节点不断发展，为了使得晶体管的面积的缩小、性能的提升尽量满足摩尔定律，即每 18 个月晶体管的密度增加一倍，同时晶体管的性能提高 15%，晶体管的结构也在不断地发展变化。如图 3.8 所示，以 SRAM 为例，其晶体管的结构从

平面晶体管(图 3.8(a))发展成 16/14nm 开始采用的鳍型场效应晶体管(图 3.8(b)),再到纳米片(Nanosheet)结构(图 3.8(c))以及叉片(Forksheets)(图 3.8(d))等结构;到了更先进的技术节点,需要采用如图 3.8(e)所示的 PMOS 与 NMOS 垂直排布的互补场效应晶体管 CFET^[3-6],此处仅列举了部分晶体管结构类型。注意,对于 PMOS 和 NMOS 的栅极需要采用不同的金属来调整其功函数,图 3.8 中只展示晶体管的大概结构。其中,Nanosheet、Forksheets 晶体管的工艺流程类似,CFET 中也包含 Nanosheet 的结构。

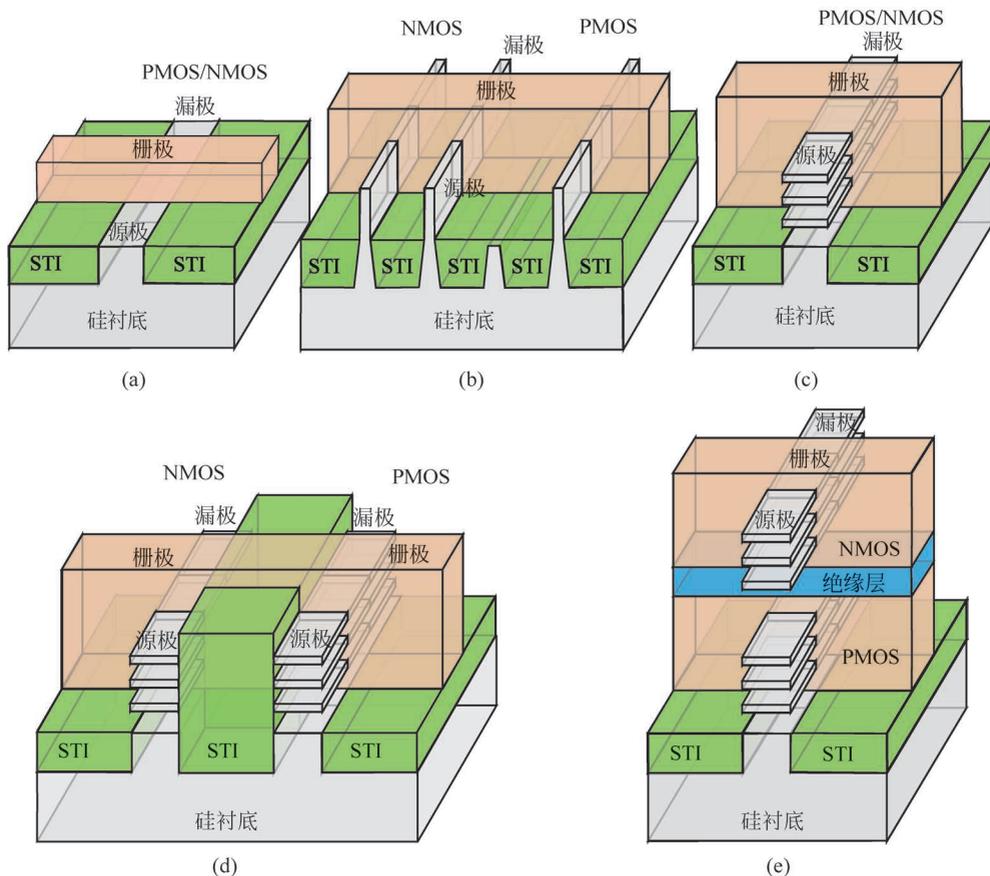


图 3.8 各种晶体管的结构示意图:(a) 平面晶体管;(b) FinFET;(c) 纳米片结构;
(d) 叉片结构;(e) CFET

接下来简单介绍其中三种常见晶体管的重点层次的工艺流程及三个技术节点关键层次的设计规则:某接近 193nm 水浸没式光刻极限的设计规则对应技术节点中的平面晶体管;14nm 技术节点的 FinFET 结构以及 3nm 技术节点的 CFET 结构(包含硅纳米片)。

3.3.2 某接近 193nm 水浸没式光刻极限的设计规则及 HKMG 平面晶体管的工艺流程简述

1. 某接近 193nm 水浸没式光刻极限的设计规则与版图设计

1) 某接近 193nm 水浸没式光刻极限的设计规则

表 3.1 是某接近 193nm 水浸没式光刻极限的前段栅极、后段金属、通孔层次的设计规则(周期,光刻最小线宽)^[7]、光刻方法、照明条件以及掩模版和光刻胶类型等信息。由于表格中

表 3.1 某接近 193nm 水浸没式光刻极限的前、后段关键层次设计规则

分段、层次信息	设计规则		光刻方法		光刻机照明条件				光掩模				光刻胶				
	层次名称	最小周期/ μm	最小线宽/ μm	光刻方法	是否有禁止周期	照明波长	照明数值孔径	照明条件种类	光源-掩模优化是否必须	偏振	掩模类型	有几层底部抗反射层	光刻胶类型	光刻胶厚度/ μm	光刻等效光酸扩散长度/ μm	光刻胶类型	显影类型
前段	栅极	117	55	单次曝光	无	193nm 水浸没	1.35	二极	否	XY	6% 相移	2	正性化学放大	90~110	5	偏高活化能	正显影
后段	金属 X	90	45	单次曝光	无	193nm 水浸没	1.35	交叉四极	否	XY	6% 相移	1	正性化学放大	90~110	5	偏低活化能	正显影
	通孔 X	110	65	单次曝光	无	193nm 水浸没	1.35	环形	否	XY	6% 相移	2	正性化学放大	90~110	5	偏低活化能	正显影

设计规则距离 193nm 水浸没式光刻机的衍射极限(约 72nm)还有一定距离,因此,均可以采用单次曝光配合合适的照明条件完成各关键光刻层次图形。另外,表格中的禁止周期是指在设计规则中被禁止掉的周期。尽管光学邻近效应修正之后,金属层次禁止周期范围的光刻工艺性能偏差(曝光能量宽裕度 EL 约为 10%),实际量产中没有真正地这一部分设计规则禁止。

2) 某接近 193nm 水浸没式光刻极限的设计规则对应技术节点中的一种 6T SRAM 的版图设计

图 3.9(a)为 3.2 节提到的 6T SRAM 电路结构,图 3.9(b)为相应的 SRAM 平面版图结构^[8-9],其中包含三层重要层次:前段有源区域(Active Area, AA)、栅极(Poly)以及中段接触孔(包含短线状的共享通孔)层次,并标注出两个反相器。红色大虚线框内为一个 SRAM 单元,图 3.9(a)和图 3.9(b)中 6 个各司其职的晶体管一一对应,有以下几点需要注意。

(1) 对于高性能的 SRAM, NMOS 晶体管的 AA 区一般大于 PMOS 晶体管的 AA 区域,以提高电流。图中只是一种 SRAM 的版图结构,实际还有很多不同性能的 SRAM 版图结构。

(2) 一个反相器中 PMOS 晶体管与 NMOS 晶体管输出(漏极)端的连接,需要通过后段金属实现,但是需要在中段设计好通孔(一个通孔,一个共享通孔)。

(3) 中段只有一层通孔,将字线(WL)、位线(BL)、电源(V_{DD})和接地(V_{SS})连接到后段的金属层次。

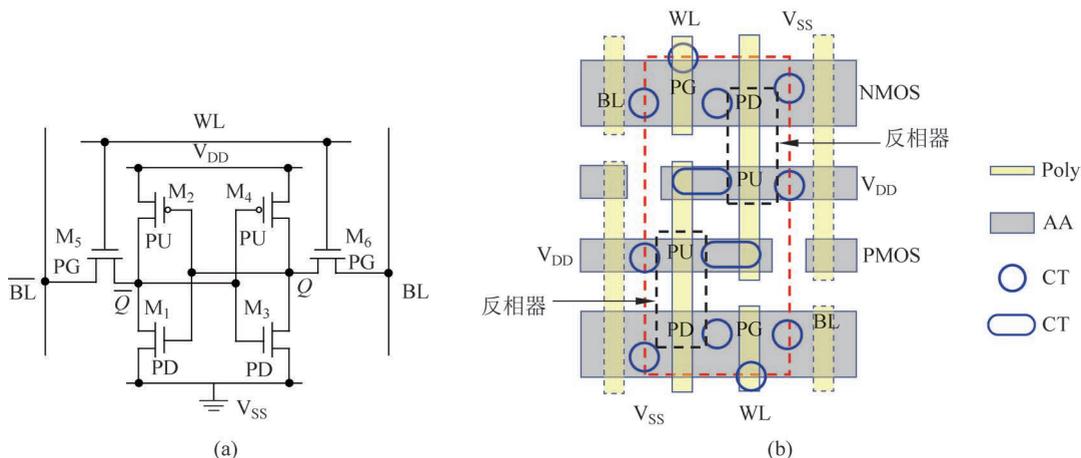


图 3.9 (a) 一种 6T SRAM 的电路和(b) 相应的平面版图结构设计示意图

2. HKMG 平面晶体管的工艺流程简述

本节以平面晶体管为例,简述 SRAM 中关键工艺流程,包括前段器件、中段和后段金属以及通孔连线形成的过程。对于平面晶体管,以高介电常数(High k)栅极介质层+金属栅极(Metal Gate, MG)(HKMG)晶体管为例,简要介绍其主要的工艺流程^[10]。如图 3.10 所示,流程主要包括:①定义有源区域;②定义双阱(Well)区域;③定义伪栅极(Dummy Poly)图形;④在栅极两侧生长间隔层侧墙 1(spacer1);⑤PMOS 和 NMOS 区域的轻掺杂漏(Light Doped Drain, LDD)离子注入工艺;⑥外延生长 PMOS 区域的 SiGe;⑦在栅极两侧生长间隔层侧墙 2(spacer2);⑧形成 PMOS 晶体管和 NMOS 晶体管的源漏区;⑨生长高介电常数栅氧和金属栅(HKMG);⑩形成中段接触孔层次(Contact Layer);⑪形成后段金属和通孔层次。

每个主要工艺流程中包含的具体内容如下。

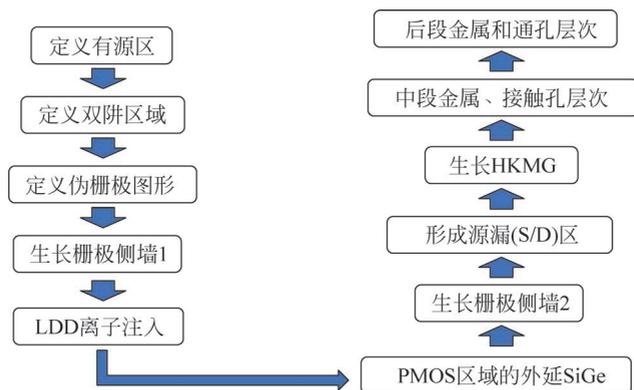


图 3.10 HKMG 平面晶体管主要工艺流程示意图

1) 定义有源区域

如图 3.11 所示,通过光刻、刻蚀、填充和 CMP 等工艺形成 NMOS 和 PMOS 晶体管的有源区。有源区之间采用浅沟道隔离(Shallow Trench Isolation, STI)技术,利用绝缘层将不同晶体管器件隔离开。晶体管性能多样,本节以 AA 区域更宽(约为 PMOS 对应 AA 区域的 2 倍)的 NMOS 晶体管为例。AA 区域更宽,可以获得更大的电流,以提高读写速率,即提高器件性能。

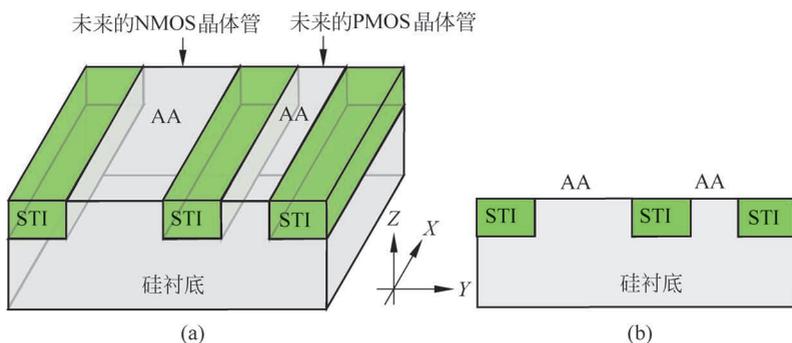


图 3.11 前段 AA 层次的定义:(a) 立体图;(b) 截面图

2) 定义双阱(Well)区域

如图 3.12 所示,定义双阱也就是定义 PMOS 和 NMOS 晶体管所在区域。

(1) 在 NMOS 区域:需要形成 P 型衬底,所以需要形成大范围的 P 型掺杂区,即 P 型阱(P Well, PW)。如图 3.12(a)所示,在形成 PW 区域时,涂覆光刻胶之后,需要通过光刻工艺将需要进行离子注入形成 PW 区域的光刻胶去除。离子注入过程中,形成 PW 区域,其他 PMOS 晶体管对应的区域有保护层(光刻胶等光刻材料)保护。

(2) 在 PMOS 区域:需要形成 N 型衬底,所以需要形成大范围的 N 型掺杂区,即 N 型阱(N Well, NW)。如图 3.12(b)所示,在形成 NW 区域时,同样需要保护层将 PW 区域保护起来,只对 PMOS 晶体管对应的区域进行离子注入。

完成离子注入工艺(去胶)之后,一般需要高温退火,不仅可以修复离子注入造成的硅晶体表面晶格损伤,还可以激活注入的离子。先进工艺中采用快速热退火,可以防止注入的离子严重扩散。

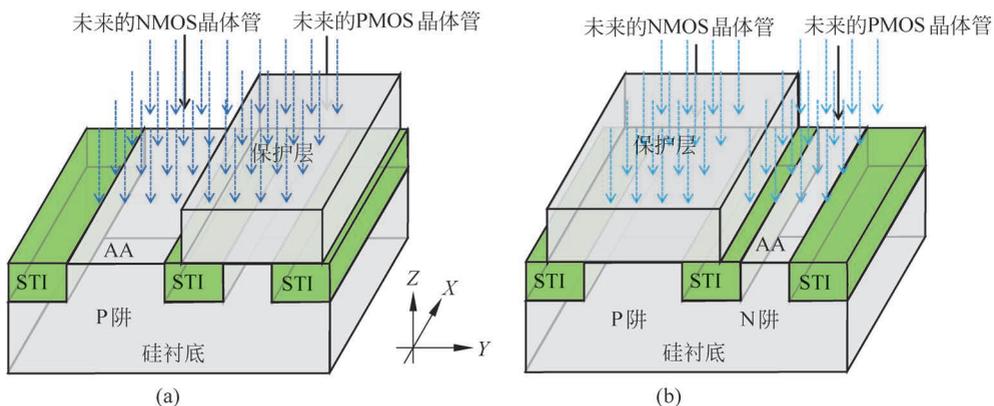


图 3.12 前段双阱层次(P Well,N Well)的定义:(a) P阱的定义;(b) N阱的定义

3) 定义伪栅极(Dummy Poly)图形

HKMG工艺,顾名思义,就是利用高介电常数的材料代替传统的栅氧材料(二氧化硅或者SiON)作为绝缘层,以改进由于晶体管微缩导致的线宽变窄,栅氧变薄引入的栅极漏电流问题。同时,用金属栅极取代多晶硅栅极以减小电阻,提高开关速度。在此之前需要生长伪栅极,如图3.13(a)所示,利用光刻和刻蚀工艺形成伪栅极图形,一般此时的伪栅极材料为无定形硅(Amorphous Silicon)。沿着 X_1 的横截面如图3.13(b)所示,栅极两侧是未来的源极(Source)和漏极(Drain)区域,此处还未进行掺杂形成源漏。另外,后续还需要有剪切层对伪栅做必要的剪切。

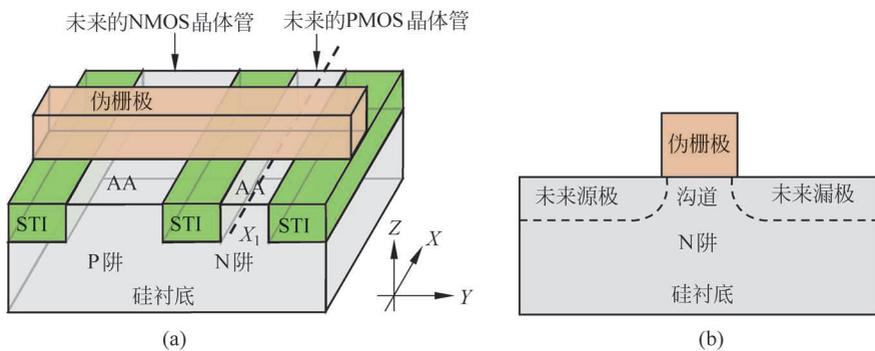


图 3.13 前段伪栅极层次定义:(a) 立体图;(b) 沿 X_1 的横截面示意图

4) 在栅极两侧生长间隔层侧墙 1(spacer1)

在进行轻掺杂漏(Light Doped Drain, LDD)离子注入工艺之前,需要先在栅极两侧生长侧墙 1(spacer1),立体图如3.14(a)所示。侧墙一般是硅化物材料,如氮化硅、氧化硅等,还可以采用低介电常数(Low k)材料。侧墙可以在LDD工艺时保护栅极,还有定位后续LDD注入位置和定义外延SiGe后的Proximity位置^[10-11]的作用,这些都会对器件性能产生很大影响。

5) PMOS和NMOS晶体管区域的轻掺杂漏(Light Doped Drain, LDD)离子注入工艺

随着栅极的宽度不断减小,栅极下方的沟道长度也不断地减小。为了有效地防止短沟道效应,需要引入轻掺杂漏工艺。可以在栅极的边界下方、沟道中靠近源漏极的附近设置一个低掺杂的漏区,该区域在源漏和沟道之间形成杂质浓度梯度,让该区域也承受部分电压,还可以减少热载流子注入效应(Hot Carrier Injection, HCI),提高器件的可靠性^[12-17]。

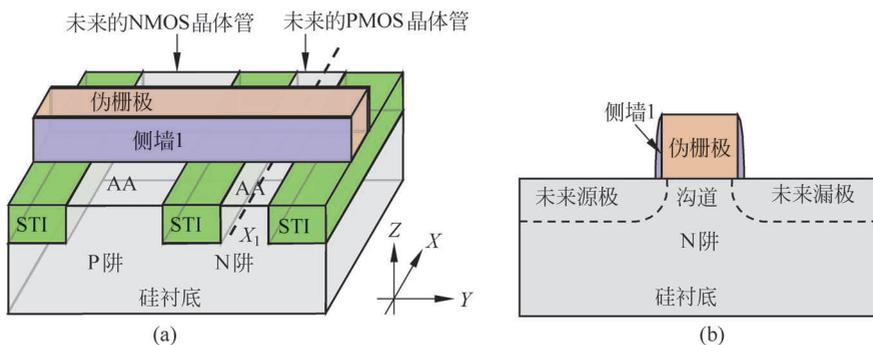


图 3.14 前段栅极两侧生长侧墙 1: (a) 立体图; (b) 沿 X_1 的横截面示意图

(1) 在 NMOS 区域,电子导电,采用的是 N 型轻掺杂漏离子注入,即 NLDD,如图 3.15(a)所示。此时,PMOS 晶体管区域需要有保护层(光刻材料或者硬掩模版等材料)保护。

(2) 在 PMOS 区域,载流子是空穴,采用的是 P 型轻掺杂漏离子注入,即 PLDD,如图 3.15(b)所示。此时,NMOS 晶体管区域需要有保护层保护。

(3) 图 3.15(c)和图 3.15(d)分别为沿着 X_2 和 X_1 的横截面,分别表示 NLDD 和 PLDD 工艺之后晶体管结构示意图。栅极两侧是未来的源极(Source)和漏极(Drain)区域,此处还未进行掺杂形成源漏。

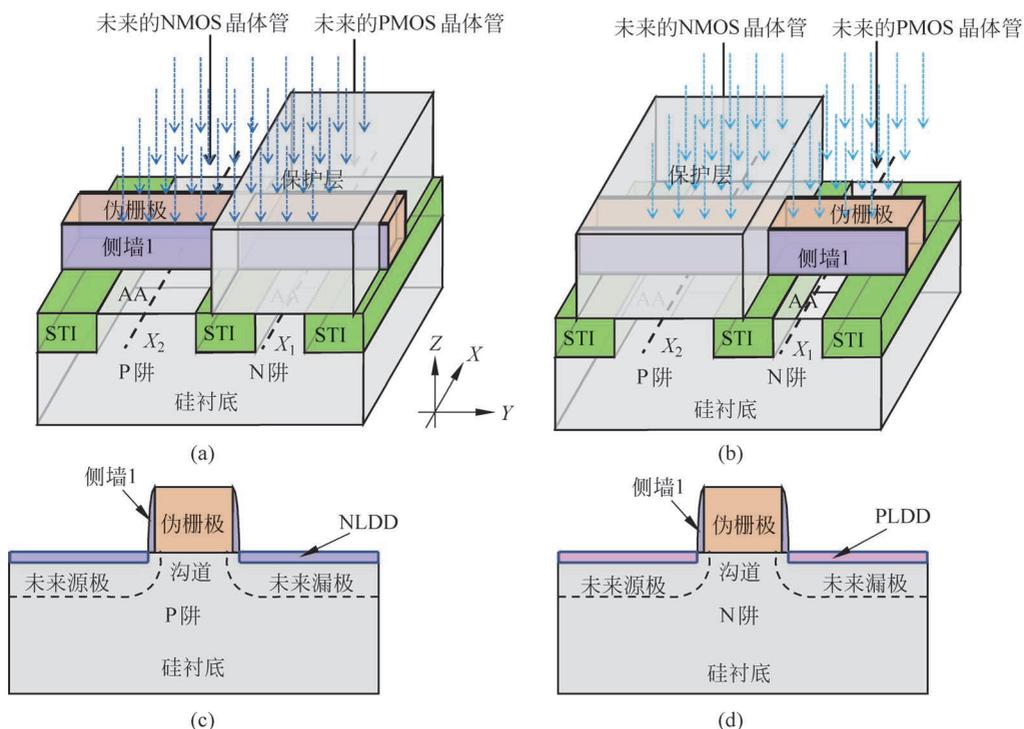


图 3.15 前段 LDD 离子注入: (a) NLDD; (b) PLDD; (c) 沿 X_2 的横截面示意图; (d) 沿 X_1 的横截面示意图

6) 生长 PMOS 区域的 SiGe

对于 NMOS 晶体管来说,其在导通时,形成的是 N 型导电通道,即导电沟道中的载流子是电子。而 PMOS 晶体管在导通时形成的是 P 型导电通道,导电沟道中的载流子是空穴。众所周知,在相同掺杂浓度下,电子的迁移率大于空穴的迁移率(2~3 倍)。NMOS 中电子迁移

率较高,电阻更小,在同掺杂的情况下,NMOS晶体管的开关速度更快。总体来说,NMOS晶体管的性能更好。

因此,需要在PMOS晶体管源漏区域经过光刻、刻蚀之后,再外延生长SiGe,通过sigma形状对沟道施加压缩的应力。图3.16(a)是生长完SiGe之后沿着如图3.15(a)所示 X_1 的横截面,这一压缩的应力可以提高空穴的迁移速度,从而提高PMOS晶体管的性能,使得PMOS晶体管的功耗、开关速度等性能尽量与同等面积的NMOS晶体管相当。一般来说,表3.1中设计规则对应的平面晶体管的NMOS晶体管中无须外延生长可以提升电子迁移速度的材料。另外,在PMOS区域生长SiGe过程中,NMOS区域(以及其他无须生长SiGe的区域)需要有硬掩模(如氮化硅)等材料保护。

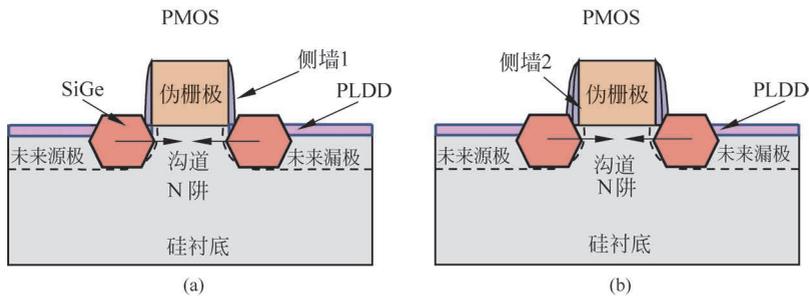


图 3.16 前段(a)PMOS晶体管外延生长SiGe; (b)栅极生长侧墙2(spacer2)示意图

7) 在栅极两侧生长间隔层侧墙2(spacer2)

在源漏离子注入工艺之前,还需要在栅极两侧继续生长侧墙(侧墙2, spacer2),如图3.16(b)所示。一般来说,这一侧墙材料仍然是低 k 材料,如氮化硅、氧化硅等,整个侧墙(spacer1 + spacer2)结构^[11]一般是多种材料的复合结构。侧墙材料可以防止金属栅极和源漏之间发生漏电,侧墙2还可以保护LDD,防止后续重掺杂的源漏离子注入破坏栅极与源漏之间的LDD低掺杂的区域。同时,侧墙2还是一种更好的精确定位源漏离子注入区的结构。

8) 形成PMOS管和NMOS晶体管的源漏区

通过离子注入工艺进行重掺杂,形成MOS晶体管的源漏区。

(1) 在NMOS晶体管的P型衬底(PW)的源漏区进行N型重掺杂,即 N^+ 工艺流程,如图3.17(a)所示,展示了 N^+ 工艺之后沿着如图3.15(a)所示 X_2 的横截面。与双阱以及LDD工艺类似,在 N^+ 工艺过程中,PMOS晶体管对应的区域需要有保护层(光刻材料或者硬掩模版等材料)保护。

(2) 在PMOS晶体管的N型衬底(NW)的源漏区进行P型重掺杂,即 P^+ 工艺流程。如图3.17(b)所示,展示了 P^+ 工艺之后沿着如图3.15(a)所示 X_1 的横截面。同样地,在 P^+ 工艺过程中,NMOS晶体管对应的区域也需要有保护层保护。

这一重掺杂过程中不能对前述栅极与源漏之间的轻掺杂LDD区域造成影响,因此栅极旁的间隔层侧墙起到了重要的隔离阻挡重掺杂的作用。同时, N^+ 和 P^+ 的离子注入工艺还可以改变通过外延工艺生长的SiGe晶格结构,从而形成中段通孔金属和SiGe的欧姆接触,减小接触电阻,增大电流,降低功耗,从而提升器件性能。

9) 生长高介电常数栅氧和金属栅(HKMG)

在生长HKMG之前,需要生长层间介电层(Inter Layer Dielectric, ILD),并通过化学机械平坦化工艺将高出栅极的ILD去掉。ILD可以在后续去除伪栅过程中,保护源漏区。接下

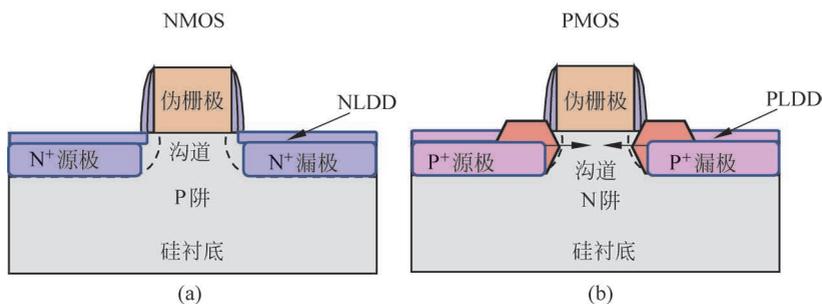


图 3.17 前段源漏(S/D)定义: (a) NMOS 晶体管 S/D 定义; (b) PMOS 晶体管 S/D 定义示意图

来,简单介绍去除伪栅以及生长 HKMG 的一般流程。

(1) 通过干法刻蚀和湿法刻蚀工艺,同时去除 NMOS 与 PMOS 晶体管的伪栅,如图 3.18(a)和图 3.18(c)所示。

(2) 同时生长界面氧化层(很薄的过渡材料)和高介电常数栅氧(例如 HfO_2)。

(3) 完成 HK 生长之后,开始生长金属栅材料。例如,先生长 PMOS 晶体管的金属栅材料(如 TiN,或者包含 TiN 在内的多层结构)。然后利用光刻工艺,保护 PMOS 晶体管,将 NMOS 晶体管中的 TiN 或者多层结构中的某些材料去除。去胶后,再生长 NMOS 晶体管的金属栅材料(如 TiAl,或者包含 TiAl 在内的多层结构)。此时,PMOS 晶体管的栅极中也会存在 NMOS 晶体管的金属栅材料(但是不会影响 PMOS 的功函数)。

(4) 统一填充低电阻金属(如 Al、W 等)^[18-22],并经过化学机械平坦化工艺磨平后,如图 3.18(b)和图 3.18(d)所示。另外,栅极功函数的调整比较复杂^[23],通过不断优化金属栅材料及材料厚度,使 NMOS 和 PMOS 的功函数达到要求。图中仅以简化的金属栅极膜层为例,说明平面晶体管中 NMOS 和 PMOS 晶体管的大概 HKMG 工艺和结构。

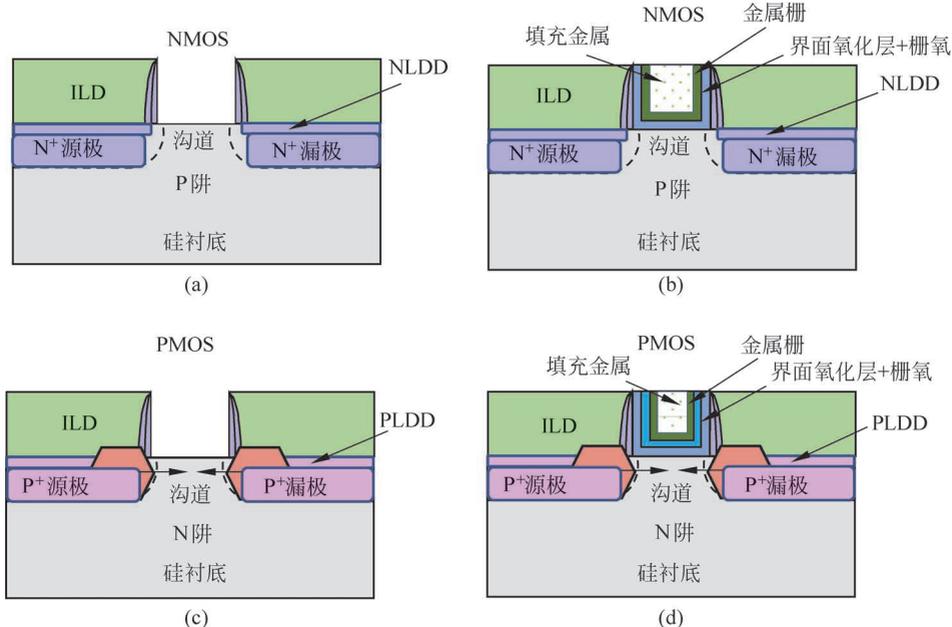


图 3.18 前段 HKMG 工艺: (a) NMOS 晶体管去除伪栅极; (b) NMOS 晶体管生长 HKMG; (c) PMOS 晶体管去除伪栅极; (d) PMOS 晶体管生长 HKMG 示意图

10) 中段接触孔层次(Contact Layer)

这一层次用来连接前段的器件(PMOS 和 NMOS)和后段的金属和通孔层次,无须按照不同类型晶体管分开完成。通过光刻、刻蚀、金属填充以及 CMP 等工艺,形成接触孔,如图 3.19 所示。其中,接触孔之间是层间介电层(ILD),接触孔材料一般为钨。

11) 后段的金属和通孔层次

从 130nm 技术节点开始,从先镀金属铝膜再刻蚀形成金属连线转变成单大马士革(Damascus)填铜工艺^[24],即先形成沟槽,再镶嵌(填充)金属铜,经过 CMP 工艺平坦化完成金属连线的埋线。而从 28nm 技术节点开始,双大马士革填铜工艺^[25]被大规模广泛应用于后段金属、通孔层次的埋线。双大马士革填铜工艺意味着同时填充两层铜线层次(一层金属和一层通孔)^[25-28],简单工艺流程如下。

- (1) 完成第一层金属(M_1)的光刻、刻蚀、金属(铜)填充和 CMP 工艺。
- (2) 完成第二层金属(M_2)的光刻和硬掩模刻蚀,然后完成第一层通孔(V_1)的光刻和部分 ILD 刻蚀,最后完成所有 ILD 刻蚀,形成 M_2 和 V_1 的最终图形。
- (3) 对 M_2 和 V_1 层统一填充金属铜,并经过 CMP 工艺去掉多余的金属。
- (4) 后续的金属 M_{x+1} 和通孔 V_x 层次($x > 1$)按照上述(2)和(3)的工艺流程完成双大马士革填铜工艺,直到完成所有金属和通孔层次的金属埋线。

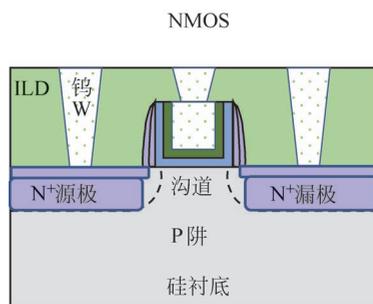


图 3.19 中(后)段接触孔层次示意图

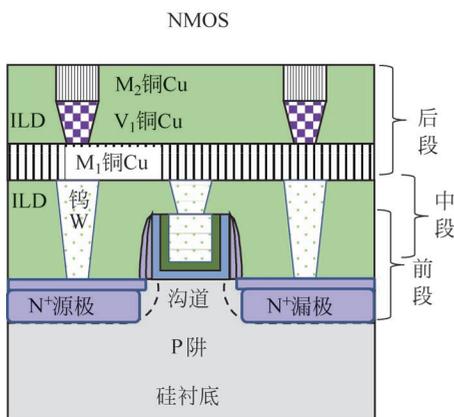


图 3.20 后段金属和通孔层次示意图

图 3.20 即为 NMOS 晶体管完成两层金属和一层通孔层次之后的沿着如图 3.15(a)所示 X_2 的横截面举例。一般来说,所有晶体管都是统一完成接触孔和后段的金属、通孔层次工艺的,无须像离子注入、HKMG 等工艺一样按照晶体管类型分开完成(即使通过多次曝光完成同一层次,填充金属也是统一完成的)。其中,中段需要将源漏以及栅极都连接到后段,因此有的通孔较深。

从图中还可以看出,后段两层金属的设计规则一般相互垂直,连通各个晶体管,而通孔负责连通两层金属。实际芯片中,会使用多层(如 6 层、8 层、10 层等)金属和通孔完成各晶体管的连接。

最后,完成顶层金属(Top Metal)、顶层通孔(Top Via)、保护层(Passivation)等工艺流程。此时,芯片工厂中的工艺流程全部结束,最后再完成封装、切割、测试等一系列后续流程,获得芯片的性能参数、可靠性以及成品率等信息。

本节通过简单的结构示意图展示了 HKMG 平面晶体管的几个关键工艺流程步骤,实际生产过程中的各工艺、膜层结构非常复杂。

3.3.3 14nm 技术节点关键层次设计规则以及 FinFET 的工艺流程简述

1. 14nm 技术节点中关键层次设计规则与版图设计

1) 14nm 技术节点中关键层次设计规则

对于 14nm 技术节点来说,设计规则开始挑战 193nm 水浸没式光刻机的衍射极限。因此,多个光刻层次需要使用多次曝光或者自对准的多重曝光来实现,如表 3.2 所示。

表 3.2 14nm 技术节点前、中、后段部分关键层设计规则^[7,29]

分段	层次名称	设计规则		光刻方法		光刻机照明条件						光掩模				光刻胶			显影类型
		最小周期/mm	最小线宽/mm	光刻方法	是否有禁止周期	照明波长	照明数值孔径	照明条件种类	光源-掩模优化是否必须	偏振	掩模类型	有几层底部抗反射层	光刻胶类型	光刻胶厚度/nm	光刻等效光酸扩散长度/mm	光刻胶类型	偏活化能		
前段	鳍	48	24	SADP	无	193nm 水浸没	1.35	二极	否	XY	6% 相移	2	正性化学放大	90~110	5	偏高活化能	正显影		
	栅极	84~90	42~45	单次曝光	有	193nm 水浸没	1.35	二极	是	XY	OMOG	2	正性化学放大	90~110	5	偏高活化能	正显影		
中段	金属 0	84~90	42~45	单次曝光 (> 80nm)	无	193nm 水浸没	1.35	二极	否	XY	6% 相移	2	正性化学放大	90~110	7	偏高活化能	负显影		
	通孔 0	64~90	32~50	LE3~LE4	无	193nm 水浸没	1.35	环形	否	XY	6% 相移	2	正性化学放大	90~110	5	偏低活化能	正显影		
后段	金属 1	64	32	LE2	无	193nm 水浸没	1.35	交叉四极	是/否	XY	6% 相移	2	正性化学放大	90~110	7	偏高活化能	负显影		
	金属 X	80	40	单次曝光	无	193nm 水浸没	1.35	二极	否	XY	OMOG	2	正性化学放大	90~110	5	偏低活化能	正显影		
	通孔 X	64~90	32~50	LE3~LE4	无	193nm 水浸没	1.35	环形	否	XY	6% 相移	2	正性化学放大	90~110	5	偏低活化能	正显影		
	通孔 X	64~90	32~50	LE3~LE4	无	193nm 水浸没	1.35	环形	否	XY	6% 相移	2	正性化学放大	90~110	7	偏高活化能	负显影		

(1) 对于前段鳍(Fin)层次来说,需要使用自对准两重图形技术配合必要的剪切层次实现小于 50nm 周期的设计规则图形。

(2) 对于后段金属层次来说,需要使用两次光刻-刻蚀(Litho-Etch Litho-Etch,LELE)^[30]实现 64nm 的、双向设计规则的最小周期。

(3) 对于中后段通孔层次来说,单次曝光可以实现的最小周期约为 90nm,而且还会有很多设计上的限制,因此最多需要使用 4 次(兼顾两个方向的图形拆分)光刻-刻蚀(Litho-Etch,LE)方法实现 64~90nm 的最小周期。其中,通孔层次周期范围与具体设计规则有关,一般是金属最小周期的 1~1.414 倍,越接近 1.414 倍,通孔周期越大,光刻工艺越容易实现。

另外,表 3.2 中的线宽更接近光刻之后线宽(尤其是单次曝光层次),实际 CDU 根据刻蚀后线宽的 10%来分配,具体可见 6.5.1 节,这会导致光刻之后的线宽均匀性需要控制在 $<\pm 10\%$ 光刻线宽目标值。

同时,中后段金属和通孔层次开始采用负显影工艺,以获得更小的光刻沟槽,减小刻蚀线宽偏置(Etch Bias),降低工艺难度,提高工艺可靠性。由表 3.2 可见,负显影光刻胶的扩散长度比正显影的偏长,这是因为负显影需要曝过光的光刻胶部分尽量不能溶于显影液,因此需要更长的等效光酸扩散长度以更充分地完成光催化反应。

2) 14nm 技术节点中一种 6T SRAM 的版图设计

图 3.21(a)为 3.2 节提到的 6T SRAM 电路示意图,图 3.21(b)为 14nm 技术节点中包含 10 Fin 的、相应的版图结构,图 3.21(c)为 14nm 技术节点中包含 8 Fin 的、相应的版图结构。版图中包含 5 层关键层次:前段 Fin、栅极(Poly)以及中段金属 0(M_0)^[31-32]层与两层接触孔层次(通孔 0 层(V_0)和接触孔($M0G$ 短线))。红色大虚线框内为一个 SRAM 单元,图 3.21(b)和图 3.21(c)中 6 个各司其职的晶体管与图 3.21(a)中的晶体管一一对应,有以下几点需要注意。

(1) 图 3.21(b)为高性能的 SRAM,NMOS 晶体管包含两根 Fin,PMOS 晶体管只包含一根 Fin;图 3.21(c)为高密度晶体管,NMOS 与 PMOS 晶体管均只包含一根 Fin。

(2) 中段包含三层金属结构(一层金属、一层接触孔和一层通孔),与 HKMG 平面晶体管类似,中段将字线(WL)、位线(BL)、电源(V_{DD})和接地(V_{SS})连接到后段的金属层次。

(3) 一个反相器中 PMOS 晶体管与 NMOS 晶体管输出(漏极)端的连接可以直接通过中段的金属 0(M_0)层次完成。

2. 14nm 技术节点中 FinFET 的工艺流程简述

本节以 14nm 技术节点高性能晶体管(NMOS 包含两根 Fin)为例,简述 SRAM 中关键工艺流程,包括前段器件、中段和后段金属以及通孔连线形成的过程。对于 14nm 技术节点的 FinFET 结构^[33],主要包含以下几个关键的工艺步骤,如图 3.22 所示:①形成 Fin;②形成伪栅极(Dummy Poly);③形成源漏(S/D)区;④生长高介电常数材料和金属栅极;⑤形成中段金属和通孔层次;⑥形成后段金属和通孔层次。

每个主要工艺流程中包含的具体内容如下。

1) 形成 Fin

(1) 定义 Fin。如图 3.23 所示,利用自对准的双重图形技术,包括光刻、刻蚀以及薄膜生长等工艺,形成 Fin 的结构。一般来说,Fin 沿 X 方向排布。

如前面所述,14nm 节点高性能的 SRAM 中,NMOS 晶体管包含两根 Fin,PMOS 晶体管包含一根 Fin。因此,完成周期性的 Fin 图形之后,需要使用 Fin 的剪切层,沿 X 方向将不需要

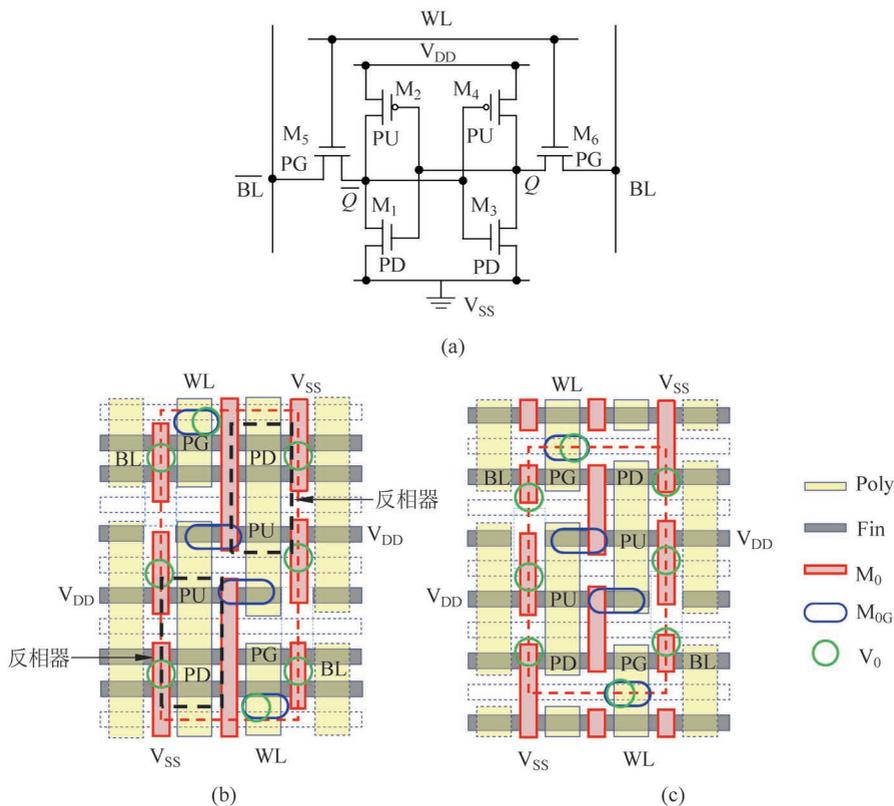


图 3.21 (a) 一种 6T SRAM 的电路; (b) 14nm 技术节点高性能 SRAM 版图设计; (c) 14nm 技术节点高密度 SRAM 版图设计示意图



图 3.22 14nm 技术节点 FinFET 结构主要工艺流程示意图

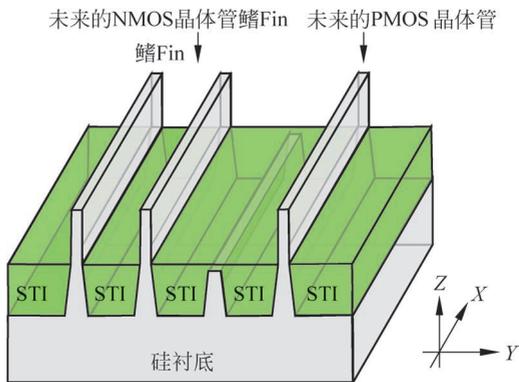


图 3.23 14nm 技术节点前段 Fin 示意图

的 Fin 切掉,这种剪切一般称为水平剪切(Horizontal Cut)。对于一根较长的 Fin,若需要从中间截断,则需要垂直剪切(Vertical Cut)层次来实现。

不同 Fin 之间需采用浅沟道隔离(STI)技术,利用绝缘层将不同晶体管器件隔离开。这里有一个 Fin 高的定义:高于 STI 区域的 Fin。

(2) 双阱(Well)区域的定义。

如图 3.24 所示,定义双阱也就是定义 PMOS 和 NMOS 晶体管所在区域。

① 在 NMOS 区域：需要形成 P 型衬底，所以需要形成大范围的 P 型掺杂区，即 P 型阱(P Well, PW)。如图 3.24(a)所示，进行双阱工艺之前，先利用 STI 填平 Fin 之间空隙。在形成 PW 区域时，涂覆光刻材料之后，需要通过光刻工艺将需要进行离子注入形成 PW 区域的光刻胶去除。离子注入过程中，形成 PW 区域，其他 PMOS 晶体管对应的区域有保护层(光刻胶等光刻材料)保护。

② 在 PMOS 区域：需要形成 N 型衬底，所以需要形成大范围的 N 型掺杂区，即 N 型阱(N Well, NW)。如图 3.24(b)所示，在形成 NW 区域时，同样需要保护层将 PW 区域保护起来，只对 PMOS 晶体管对应的区域进行离子注入。

双阱工艺完成之后，还需要将填平所用的 STI 去掉，露出一定高度的 Fin。

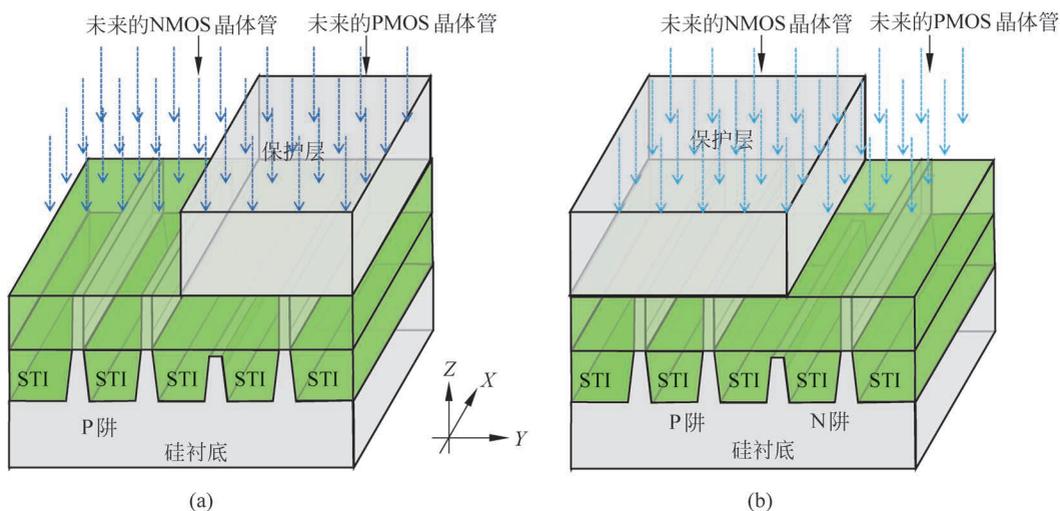


图 3.24 14nm 技术节点前段双阱层次的定义：(a) P 阱的定义；(b) N 阱的定义

2) 形成伪栅极

(1) 伪栅极(Dummy Poly)图形定义。

对于 14nm 技术节点来说，栅极设计规则还未挑战 193nm 水浸没式的衍射极限，仍然可以采用单次曝光完成。对于周期较小的设计规则，如 78nm 周期(与器件无关的栅极周期)，需要使用强偶极照明条件。经过光刻和刻蚀工艺之后，形成伪栅极图形，如图 3.25 所示，一般此时的伪栅极材料为无定形硅，且方向与 Fin 垂直，即沿着 Y 方向排布。另外，后续还需要有剪切层对伪栅做必要的剪切。

(2) 如 3.3.2 节 HKMG 平面晶体管工艺流程所述，在进行轻掺杂漏(Light Doped Drain, LDD)离子注入工艺之前，需要先在栅极两侧生长侧墙 1(spacer1)。

(3) PMOS 和 NMOS 晶体管区域的轻掺杂漏离子注入工艺。

如 3.3.2 节 HKMG 平面晶体管工艺流程所述，需要引入轻掺杂漏工艺。

① 在 NMOS 区域，电子导电，采用 N 型轻掺杂漏离子注入，即 NLDD。此时，PMOS 晶体管区域需要有保护层(光刻材料或者硬掩模等材料)保护。

② 在 PMOS 区域，载流子是空穴，采用 P 型轻掺杂漏离子注入，即 PLDD。此时，NMOS 晶体管区域需要有保护层保护。

3) 形成源漏(S/D)区

(1) 与 3.3.2 节平面晶体管类似，源漏离子注入工艺之前，还需要再沿着栅极方向(与 Fin

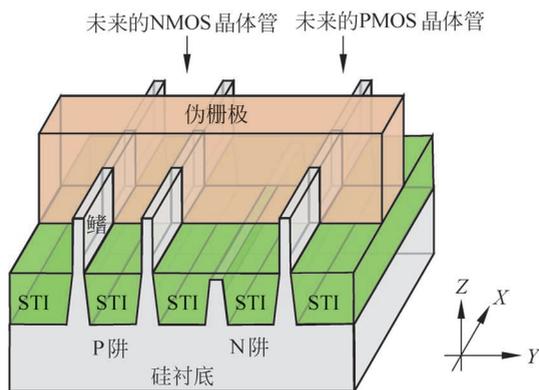


图 3.25 14nm 技术节点前段伪栅极层次定义示意图

垂直方向),在栅极两侧生长间隔层侧墙 2(spacer2)。

(2) 外延生长 PMOS 区域的 SiGe 和 NMOS 区域的 SiP,分别对 PMOS 和 NMOS 区域进行掺杂形成源漏区。

由 3.3.2 节 HKMG 平面晶体管工艺流程所述,需要进行源漏区域的外延生长。

① 对于 PMOS 晶体管源漏区域,需要生长 SiGe,对沟道施加压缩的应力。在生长 SiGe 之前,需要通过光刻、刻蚀等工艺,刻蚀掉 PMOS 源漏区的 Fin 图形。此时,NMOS 区域和 PMOS 栅极区域需要被硬掩模与光刻胶等材料保护。随后,在 PMOS 源漏区域,沿着衬底上的硅外延生长 SiGe,外延生长 SiGe 时,其他区域被硬掩模(如氮化硅)等材料保护。

② 对于 NMOS 晶体管源漏区域,需要生长 SiP,对沟道施加拉伸的应力。同样,在生长 SiP 之前,需要通过光刻、刻蚀等工艺,刻蚀掉 NMOS 源漏区的 Fin 图形。此时,PMOS 区域和 NMOS 栅极区域需要被硬掩模与光刻胶等材料保护。随后,在 NMOS 源漏区域,沿着衬底上的硅外延生长 SiP,外延生长 SiP 时,其他区域被硬掩模(如氮化硅)等材料保护。最后结果如图 3.26 所示,其中,栅极两侧有侧墙(spacer1+spacer2,多层复合结构)。

完成外延生长之后,还需要分别完成 NMOS 和 PMOS 晶体管的源漏离子注入工艺,以形成 MOS 晶体管的源漏区域。完成其中一种晶体管的源漏离子注入工艺时,另一种晶体管对应区域需要保护层保护。

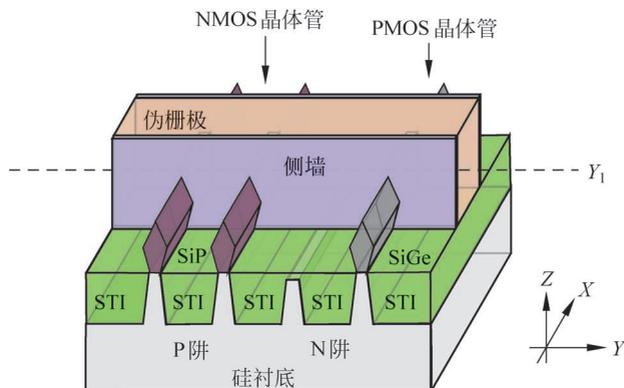


图 3.26 14nm 技术节点前段侧墙、外延工艺完成之后示意图

4) 生长高介电常数栅氧和金属栅(HKMG)

与 3.3.2 节 HKMG 平面晶体管类似,在生长 HKMG 之前,也需要生长 ILD,并通过化学

机械平坦化工艺将高出栅极的 ILD 去掉。ILD 可以在后续去除伪栅过程中,保护源漏区。接下来,简单介绍去除伪栅以及生长 HKMG 的一般流程。

① 通过干法刻蚀和湿法刻蚀工艺,同时去除 NMOS 与 PMOS 晶体管的伪栅。

② 同时生长界面氧化层(很薄的过渡材料)和高介电常数栅氧(如 HfO_2)。

③ 完成 HK 生长之后,开始生长金属栅材料。例如,先生长 PMOS 晶体管的金属栅材料(如 TiN,或者包含 TiN 在内的多层结构)。然后利用光刻材料,保护 PMOS 晶体管,将 NMOS 晶体管中的 TiN 或者多层结构中的某些材料去除。去胶后,再生长 NMOS 晶体管的金属栅材料(如 TiAl,或者包含 TiAl 在内的多层结构)。此时,PMOS 晶体管的栅极中也会存在 NMOS 晶体管的金属栅材料(但是不会影响 PMOS 的功函数)。

④ 统一填充低电阻金属(如 W)^[34-35],并经过化学机械平坦化工艺磨平。

沿着如图 3.26 所示,与栅极平行且经过栅极中心的切线 Y_1 ,经过 HKMG 工艺之后,其截面如图 3.27 所示。图中以简化的金属栅极膜层为例,说明 FinFET 晶体管中 NMOS 和 PMOS 晶体管的大概 HKMG 工艺和结构。

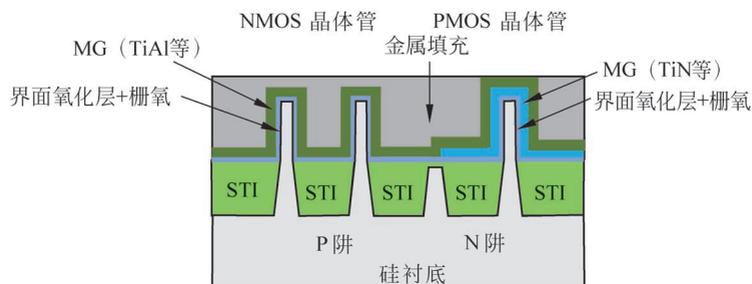


图 3.27 14nm 技术节点前段 HKMG 工艺结束之后示意图

接下来,通过中段的金属层次、接触孔、通孔以及后段的金属和通孔层次将前段器件导出,具体可参考 3.3.4 节的 CFET 中后段结构。尽管 3nm 节点与 14nm 节点在中后段所用金属材料不同,但是分段规则和各段的功能有相通性,本节不再赘述。

本节通过简单的结构示意图展示了 14nm FinFET SRAM 中的几个关键工艺流程步骤,实际生产过程中的各工艺、膜层结构非常复杂。

3.3.4 3nm 关键层次设计规则以及 CFET 的工艺流程简述

1. 3nm 技术节点中关键层次设计规则

如前所述,从 16/14nm 技术节点开始引入鳍型晶体管(FinFET),而到了更加先进的技术节点,例如 3nm 技术节点,可以开始采用 CFET 结构,以大大减小 SRAM 面积。本节以 3nm CFET 结构为例,简述 3nm 节点关键层次设计规则、重要步骤的工艺流程。表 3.3 是 3nm 节点前、中、后段关键层次的主要设计规则、曝光条件以及光刻材料等信息。

(1) 对于前段鳍(纳米板)图形,需要使用自对准四重图形技术配合必要的剪切层次实现约 24nm 周期的设计规则图形。对于前段栅极图形,需要使用自对准两重图形技术配合必要的剪切层次实现小于 50nm 周期的设计规则图形。一般来说,当单次曝光周期大于或等于 38nm,小于 76nm 时,即可采用两次(包括自对准)193nm 水浸没式光刻工艺和刻蚀工艺(LELE,SALELE)或者 193nm 水浸没式光刻工艺和自对准双重图形技术(SADP)的方法完成图形。

表 3.3 3nm 技术节点前、中、后段关键层次设计规则^[29,36]

分段、层次信息	设计规则		光刻设计规则 (一次光刻)				光刻机照明条件				光掩模				光刻胶				
	层次名称	最小周期/mm	最小线宽/mm	光刻方法	是否有禁止周期	照明波长	照明数值孔径	照明条件种类	光源-掩模优化是否必须	偏振	掩模类型	有几层底部抗反射层	光刻胶类型	光刻胶厚度/mm	光刻等效光酸扩散长度/mm	光刻胶类型	显影类型		
前段	鳍(纳米板)	24	12	SAQP	无	193nm 水浸没	1.35	二极	否	XY	6%相移 2	正性化学放大	90~110	5	偏高能	正显影			
	BPR 层	120	30	单次曝光	无	193nm 水浸没	1.35	交叉四极	否	XY	6%相移 2	正性化学放大	90~110	7	偏高能	负显影			
	栅极	48	24	SADP	无	193nm 水浸没	1.35	弱二极	否	XY	6%相移 2	正性化学放大	90~110	5	偏高能	正显影			
中段	金属 0	48	24	单次曝光	无	0.33NA EUV	0.33	四极	否	无	无	正性化学放大	40	4	偏低能	正显影			
	通孔 0	24~36	12~18	LE2~LE3	无	0.33NA EUV	0.33	环形	否	无	无	正性化学放大	40	4	偏低能	正显影			
后段	金属 X	24	12	SALELE	无	0.33NA EUV	0.33	四极	否	无	无	正性化学放大	40	4	偏低能	正显影			
	通孔 X	24~36	12~18	LE2~LE3	无	0.33NA EUV	0.33	环形	否	无	无	正性化学放大	40	4	偏低能	正显影			

(2) 在 5nm 技术节点以下的先进工艺中,前段工艺流程会引入埋入式电源线(Buried Power Rail,BPR)^[6]层次,以减轻中后段金属布线的压力。

(3) 一般来说,从 5nm 节点开始,前段的剪切层次和中后段的金属和通孔层次最好采用 EUV 光刻工艺。由于存在光子吸收随机效应,实际采用 0.33 NA EUV 光刻工艺的周期距离衍射极限(约 22nm)较远:对于金属层次,单次 EUV 曝光的最小周期约为 36nm;对于通孔层次,单次 EUV 曝光的最小周期约为 48nm。

因此,如表 3.3 所示,中段金属可以采用 0.33 NA EUV 单次曝光,中段通孔需要根据实际设计规则选择采用几次(2~3 次)EUV 光刻-刻蚀工艺。

(4) 后段金属需要采用自对准的两次 EUV 光刻-刻蚀工艺(EUV SALELE),后段通孔层次也需要根据实际设计规则选择采用几次(2~3 次)EUV 光刻-刻蚀工艺。

注意,表 3.3 中的设计规则是基于一种 3nm CFET SRAM 结构来定义的。实际每家公司针对 3nm 技术节点会有不同的结构设计,例如,仍然保持 FinFET 结构,即使采用 CFET 结构,SRAM 的具体设计也会跟本节的有所差别。

2. CFET 与 FinFET 的版图区别

图 3.28 展示了两种 6T SRAM 的平面版图结构,包含几个关键层次:前段 Fin,栅极,中段的金属 0 层(M_0)、通孔 0 层(V_0)和接触孔(M_0G)层次。到了先进技术节点,若想提高晶体管密度,SRAM 中 PMOS 晶体管与 NMOS 晶体管中可以只包含一根 Fin。接下来分别介绍这两种 SRAM 的具体结构。

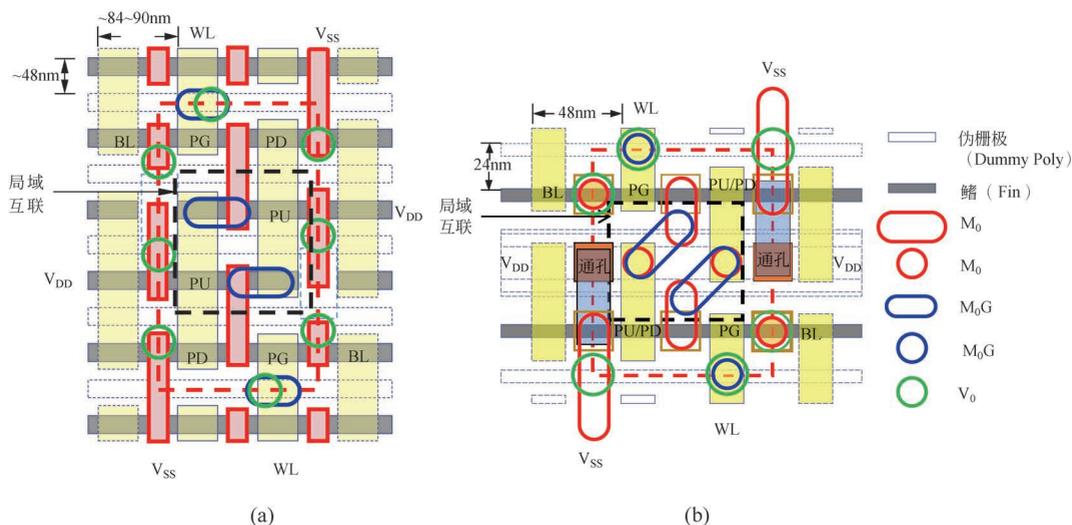


图 3.28 两种 6T SRAM 版图示意图:(a)一种高密度的 14nm FinFET 结构;(b)一种高密度的 3nm CFET 结构

图 3.28(a)是包含 6 个晶体管的、高密度的 14nm FinFET 结构,红色大虚线框中代表一个 SRAM 基本单元,由图中可知:

(1) SRAM 包含两个 PG 晶体管,其栅极会经过中段的通孔层次连通到后段的金属层次,此金属层次作为字线(WL)。

(2) SRAM 包含左下和右上两个反相器。两个反相器各自的输出作为对方的输入(输入到栅极),通过中段局域互联层次(接触孔)实现。

(3) 每个反相器中包含一个 NMOS(PD)和一个 PMOS(PU)晶体管,一个反相器中两个

晶体管输出(漏极)需要同时输入另外一个反相器的栅极,因此需要通过中段金属 0 层次连通一个反相器中两个晶体管的输出端。

(4) 两个反相器中的两个 PMOS(PU)晶体管源极需要接电源,通过中段的通孔 0 层次、中段的金属 0 层次将两个 PU 晶体管的源极连接到后段的金属层次,此金属层次接电源(V_{DD})。

(5) 两个反相器中的两个 NMOS(PD)晶体管源极需要接地,通过中段的通孔 0 层次、中段的金属 0 层次将两个 PD 晶体管的源极连接到后段的金属层次,此金属层次接地(V_{SS})。

(6) 两个位线分别连接两个 PG 晶体管的源极,通过中段的通孔 0 层次、中段的金属 0 层次将两个 PG 晶体管的源极连接到后段的金属层次,此金属层次作为位线(BL)。

(7) 另外,由于晶体管之间需要隔离,所以需要利用水平剪切层将 Fin 隔一根切掉一根。栅极与金属 0 层次需要形成的线端-线端尺寸较小,193nm 水浸没式单次光刻工艺无法实现,也需要通过剪切实现。因此,图中栅极与金属 0 层次的线端为方形。

图 3.28(b)为一种高密度的、3nm CFET 6T SRAM,红色大虚线框中代表一个 SRAM 基本单元,具体晶体管结构均与 14nm FinFET 的类似,区别如下。

(1) 为了继续缩小 SRAM 面积,从 3nm 技术节点可以开始采用基于硅纳米板的 CFET 结构,即一个反相器中的 PMOS 和 NMOS 在垂直方向叠加放置,如图中“PU/PD”所示。

(2) 传统的互联方式为横平竖直的设计规则,图中为了尽量减小 SRAM 面积,同时节省中段局域互联的掩模版,图中尝试了一种 45° 局域互联的方式。

(3) 同样地,由于晶体管之间需要隔离,所以需要利用水平剪切层将 Fin 切掉,本设计只用 5 根 Fin,且处于局域互联区域的 Fin 被切掉两根,其余切掉一根。栅极需要形成的线端-线端尺寸较小,193nm 水浸没式光刻工艺无法实现,需要通过剪切(EUV 光刻)实现,因此图中栅极线端为方形。而金属 0 层次的小线端-线端尺寸可以通过单次极紫外光刻工艺来实现,因此图中金属 0 线端为圆形(光学邻近效应)。

(4) 另外,对于 CFET 结构来说,一般 PMOS 在下方,NMOS 堆叠在 PMOS 上方。这是因为,一个 6T SRAM 中有 4 个 NMOS,将较多的 NMOS 放在上方来制作可以简化工艺。同时,NMOS 晶体管的性能比 PMOS 的性能要好,将 NMOS 晶体管置于上方可以防止处于上方的 PMOS 晶体管在高温($600\sim 700^\circ\text{C}$)^[37]外延工艺过程中损伤处于下方的 NMOS 晶体管。

3. 3nm 技术节点中 CFET 的工艺流程简述

本节以 3nm 技术节点为例,简述采用 CFET 结构时 SRAM 的基本工艺流程^[38]并展示重要工艺步骤。同时以 SRAM 版图平面和部分截面图为例,简述一个完整 SRAM 形成的过程。如图 3.29 所示,流程主要包括:①形成 Fin,埋入式电源线(BPR);②形成伪栅极(Dummy Poly),源漏(S/D)区,其中,PMOS 使用 SiGe,NMOS 使用 SiP,生长高介电常数材料和金属栅极;③形成中段金属和通孔层次;④形成后段金属和通孔层次。

每个主要工艺流程中包含的具体内容如下。

1) 形成 Fin 和 BPR

对于 3nm CFET 来说,采用的是硅纳米板结构,因此需要先在衬底上生长硅(Si)与锗硅

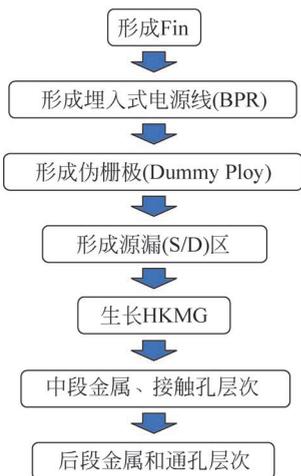


图 3.29 3nm 技术节点中采用 CFET 结构时 SRAM 的主要工艺流程示意图

(SiGe)的多层膜,本章以 PMOS 和 NMOS 各三层硅纳米板为例。图 3.30 为完成前段鳍和 BPR 之后的平面图和一个 6T SRAM 单元(红色虚线框)内 4 处截面示意图: X, Y, Y_2, Y_3 。具体内容如下。

(1) 对于鳍层次来说,利用自对准的四重图形技术(SAQP),包括光刻、刻蚀以及薄膜生长等工艺,形成 Fin 的结构。不同 Fin 之间需采用浅沟道隔离(STI)技术,利用绝缘层将不同晶体管器件隔离开。

(2) 图 3.30 中, X 是一个 SRAM 单元内沿着 Fin 的方向(一般为 X 方向)的横截面,可以看出垂直叠加的 PMOS 和 NMOS 中靠外延生长的硅纳米板,其中两种 MOS 晶体管中间的是靠外延生长的、高掺杂 Ge 的 SiGe,后续会被绝缘层(例如氮化硅)代替。而硅纳米板之间是 Ge 掺杂量较低的 SiGe,后续会被高介电常数(HK)栅氧以及金属栅代替^[38]。

(3) Y, Y_2, Y_3 是一个 SRAM 单元内沿着垂直 Fin 的方向或者平行于栅极方向的横截面,由于现在还没有栅极和源漏,所以三个横截面是相同的结构:除了 Fin,还有处于局域互联位置下方的埋入式电源线——BPR。在这一 SRAM 设计中,BPR 作为电源 V_{DD} ,经过光刻、刻蚀、填充金属以及回刻等工艺实现。在 BPR 填充金属过程中,需要将 Fin 保护起来。

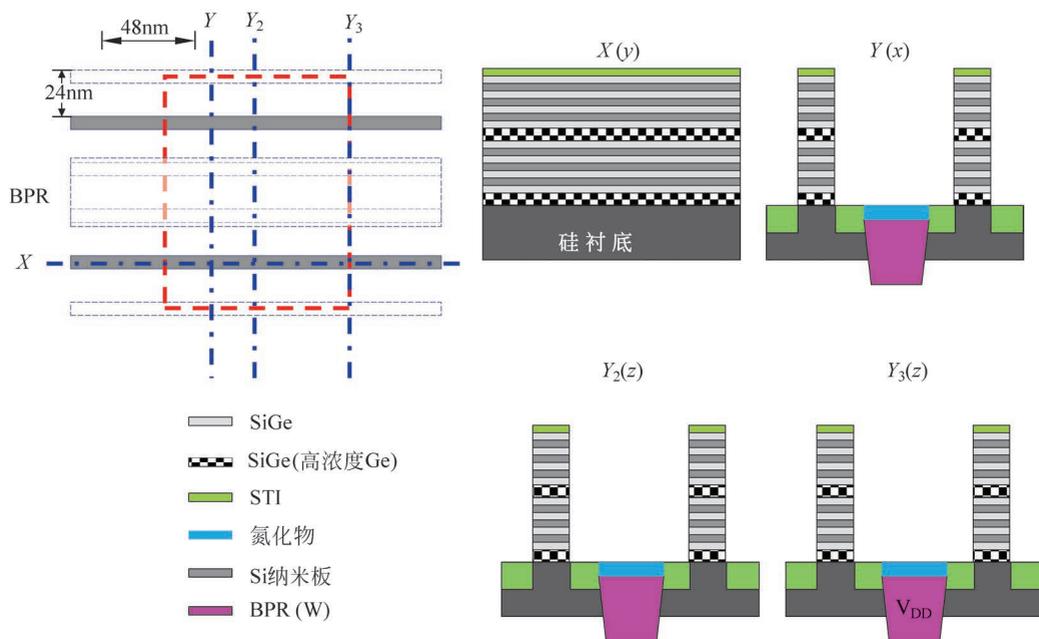


图 3.30 3nm 技术节点前段 Fin 和 BPR 工艺完成后的结构示意图

2) 形成伪栅极、源漏区和 HKMG

图 3.31 为完成前段伪栅极,伪栅极间隔层(Spacer),定义源漏区域,内部间隔层(Inner Spacer),源漏外延层生长(EPI axial (EPI) Growth)、离子注入,PMOS 的通孔(连接电源 V_{DD})以及 HKMG 等一系列工艺之后的平面图和一个 6T SRAM 单元(红色虚线框)内 4 处截面示意图: X, Y, Y_2, Y_3 。具体内容如下。

(1) 对于伪栅极层次来说,利用自对准的双重图形技术(SADP),包括光刻、刻蚀以及薄膜生长等工艺,形成伪栅极的结构。生长栅极间隔层,同时也会在掏空了的衬底与晶体管(硅纳米板)之间以及垂直晶体管之间间隙生长绝缘层(如氮化硅材料)。

(2) 如上所述, X 是一个 SRAM 单元内沿着鳍的方向(一般为 X 方向)的横截面,会经过

三处源漏(S/D)区域和两处栅极。在进行源漏区域定义(光刻、刻蚀)时,由于硅纳米板之间是低掺杂 Ge 的 SiGe,刻蚀速率高于硅,会被部分刻蚀,造成硅纳米板与 SiGe 之间有空隙,需要先生长 Inner Spacer 作为绝缘层,再完成源漏区的外延。

(3) Y 是沿着与 X 切线垂直方向,沿着栅极方向并经过两根鳍的横截面。可以看出,在生长 HKMG 之前,需要生长 ILD,并通过化学机械平坦化工艺将高出栅极的 ILD 去掉。ILD 可以在后续去除伪栅过程中,保护源漏区。再将伪栅掏空,此时栅极覆盖下鳍中硅纳米板中间的 SiGe 也会被掏空。

同时生长界面氧化层(很薄的过渡材料)和高介电常数栅氧(如 HfO_2)。完成 HK 生长之后,开始生长金属栅材料。例如,先生长 PMOS 晶体管的金属栅材料。然后利用光刻材料(如 Spin On Carbon, SOC)保护下方的 PMOS 晶体管,将 NMOS 晶体管中的金属栅材料去除。去除光刻材料之后,再生长 NMOS 晶体管的金属栅材料。此时,PMOS 晶体管的栅极中也会存在 NMOS 晶体管的金属栅材料(但是不会影响 PMOS 的功函数)。最后,也需要利用低电阻金属(如 W)进行填平^[39],如图 3.31 中 Y 截面所示(栅极已经完成必要的剪切并填充了绝缘层 ILD)。

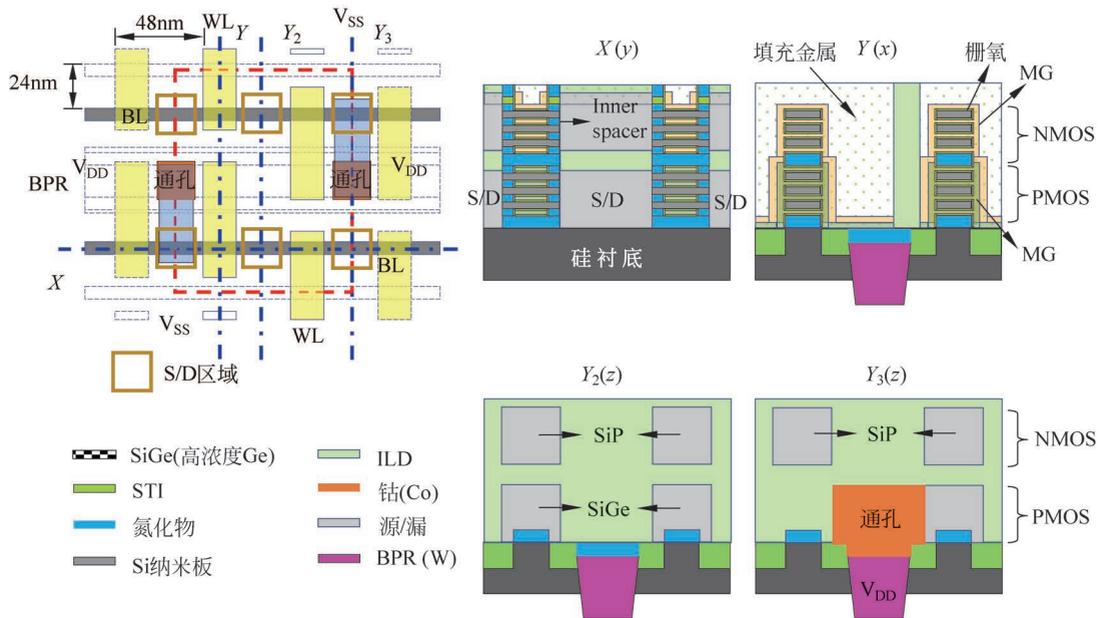


图 3.31 3nm 技术节点前段伪栅极、源漏及 HKMG 工艺完成后的结构示意图

(4) Y_2 也是沿着与 X 切线的垂直方向,同时还横切两个反相器的漏极(输出端),因此可以看到两个反相器中处于下方的 PMOS 的外延——SiGe,以及处于上方的 NMOS 的外延——SiP,PMOS 与 NMOS 晶体管的外延层之间通过绝缘层隔离。需要注意的是,由于垂直排布的 PMOS 与 NMOS 中外延材料不相同,所以需要分批进行:先外延 PMOS 晶体管的 SiGe,回刻后再外延 NMOS 晶体管的 SiP。

(5) Y_3 同样是沿着与 X 切线的垂直方向,除了会横切后续工艺流程中从后段连接出去(此处还未显示)的接地(V_{SS})和位线(BL)对应的源极,还会横切 PMOS 的源极以及将此源极连接电源(BPR/ V_{DD})的通孔。

(6) 可以看出,由于 Y, Y_2, Y_3 的方向均与鳍以及 BPR 的方向垂直,因此三者的横截面都

会显示 BPR。

3) 形成中段金属和通孔层次

图 3.32 为完成中段金属和通孔层次(金属 0 层,接触孔以及通孔 0 层次)工艺之后的平面图和一个 6T SRAM 单元(红框)内 4 处截面示意图: X, Y, Y_2, Y_3 。具体内容如下。

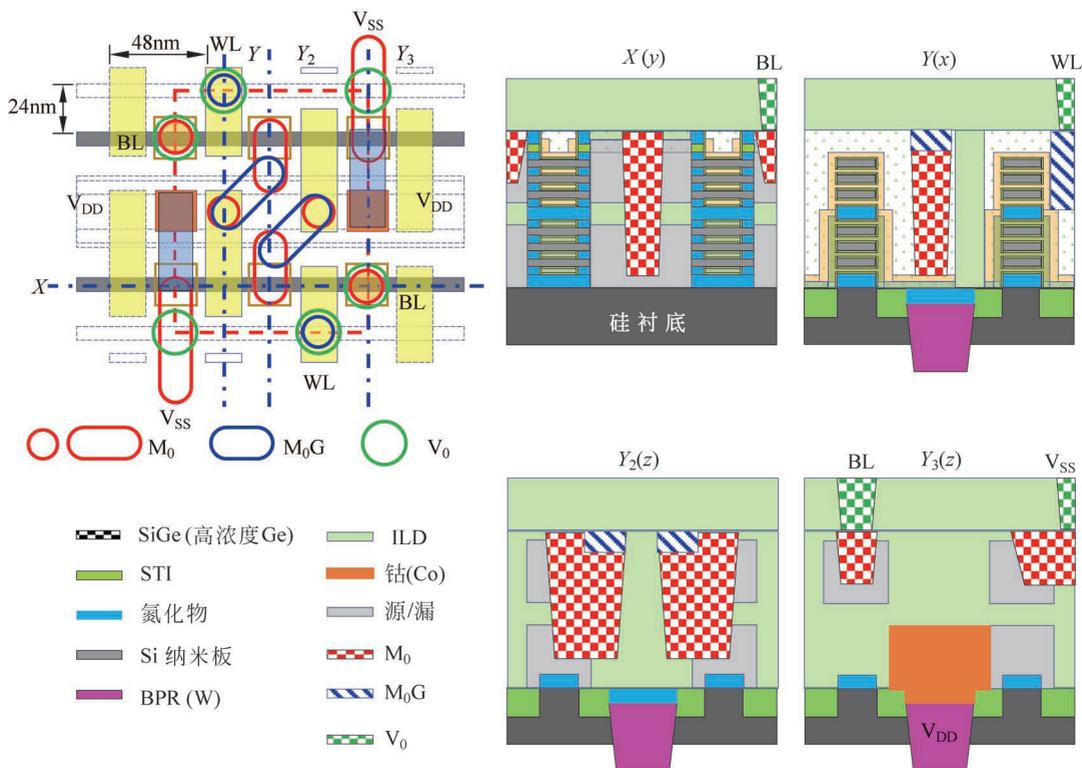


图 3.32 3nm 技术节点中段金属(M_0)、接触孔(M_0G)以及通孔 0 层次(V_0)工艺完成后的结构示意图

(1) X 是沿着鳍的方向(一般为 X 方向)的横截面, X 切线从左到右依次经过三处 M_0 。其中,只有中间一处 M_0 需要连通上下两层 MOS 晶体管的输出(漏极),因此深度更深。这是因为此处作为 SRAM 中其中一个反相器(包含一个 PMOS 和一个 NMOS 的处于平面 SRAM 左下方的反相器)的漏极需要通过局域互联与另外一个反相器(处于 SRAM 右上方的反相器)的栅极相连。

在这个 SRAM 中,沿着 X 切线位置最右端是一个位线(BL)的通孔。从 X 切线的截面图中也可以看到这个通孔 0(V_0)层次,最终位线通过金属 1(M_1)层次连接出去,见图 3.33。

(2) Y 是沿着与 X 切线垂直方向,沿着栅极、从下往上并经过两根鳍的横截面。从 Y 截面图中(从左到右)可以看出,经过的第一个金属 0 层次以及局域互联层(M_0G)是为了将这个处于平面 SRAM 左下方的反相器的栅极(输入)与处于右上方反相器的漏极(输出)连接在一起;经过的第二处局域互联(M_0G)以及通孔 0 层是为了将此 SRAM 的字线(WL)连接出去,字线最终通过金属 2(M_2)层次连接出去,见图 3.33。

(3) 在中段工艺流程中,可以看到 Y_2 截面中也新增加了金属 0 层次和局域互联层次(M_0G),其中,金属 0 贯通 PMOS 和 NMOS,是为了连接一组反相器中的两个 MOS 晶体管的漏极(输出)。如上文所述,再通过金属 0 上方的局域互联层次(M_0G)将一组反相器的漏极(输出)与另外一组反相器的栅极(输入)相连。

(4) 在中段工艺流程中,从平面图中沿着栅极的方向从下往上横切,最终形成 Y_3 从左往右的截面图,可以看到 Y_3 的截面图中新增加了两组金属 (M_0) 和通孔 (V_0) 层次,每组 M_0 和 V_0 分别连接到 SRAM 中上层的两个 NMOS 晶体管。其中, Y_3 截面图中从左往右的第一组 M_0 和 V_0 是将位线(上文中提到的 X 切线中的位线)连接到上层金属 (M_1),第二组 M_0 和 V_0 是将接地线 (V_{SS}) 连接到上层的金属 (M_2)。

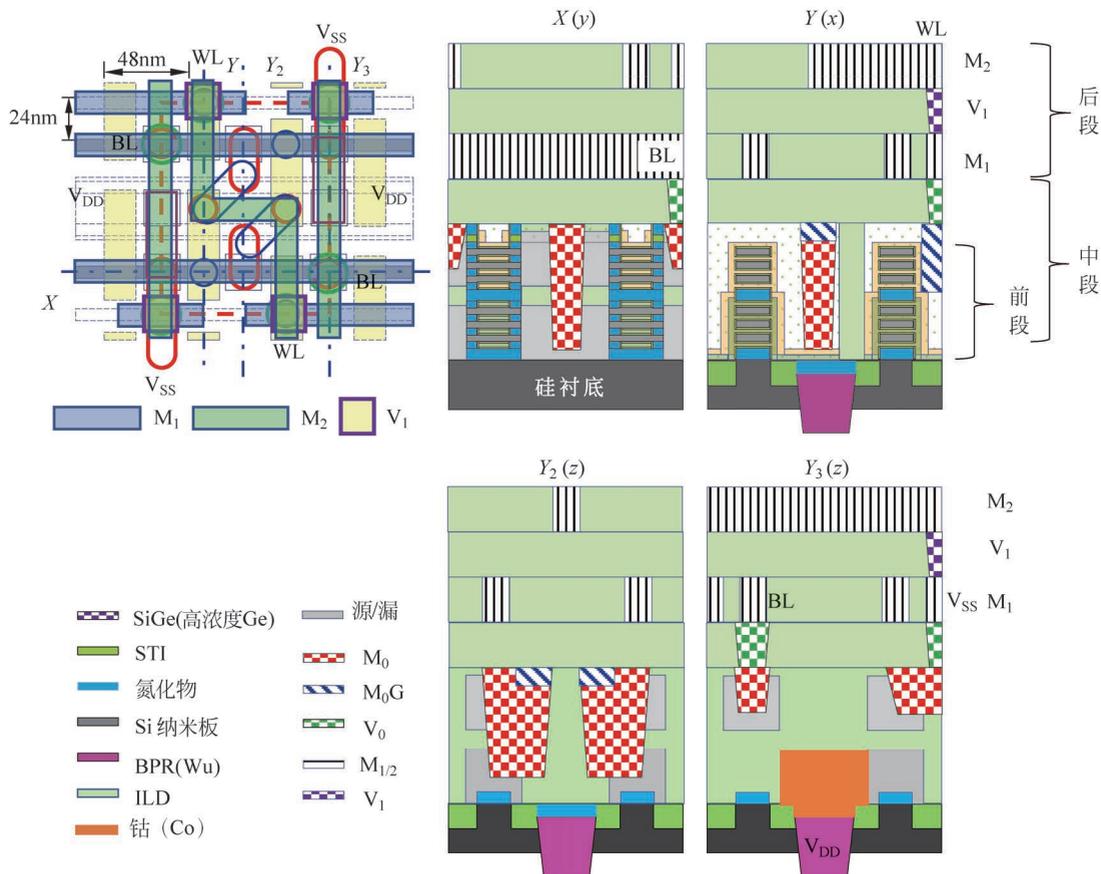


图 3.33 3nm 技术节点后段部分金属和通孔层次工艺完成后的结构示意图

4) 形成后段金属和通孔层次

经过上述前段器件以及中段的互联工艺之后,接下来是后段的金属和通孔层次,例如, $M_1, V_1, M_2, V_2, M_3, V_3, \dots$ 工艺流程。图 3.33 即为完成后段金属 1、2 和通孔 1 层次之后的平面图和一个 6T SRAM 单元(红框内) 4 处截面示意图: X, Y, Y_2, Y_3 。具体内容如下。

(1) 其中,位线(BL)、接地(V_{SS})通过 M_1 连接出去,字线通过 M_2 连接出去。

(2) 整个工艺流程分为前段(器件)、中段以及后段。其中,中段需要将源漏以及栅极都连接到后段,因此有的层次(例如 M_0) 深度不一致。

(3) 最后还会有常规的平面晶体管以及 FinFET 中类似的顶层金属(Top Metal)、顶层通孔(Top Via)以及保护层(Passivation)等工艺流程。此时,芯片工厂中的工艺流程全部结束,最后再完成封装、切割、测试等一系列后续流程,获得芯片的性能参数、可靠性以及成品率等信息。

本节通过简单的结构示意图展示了 3nm CFET SRAM 中的几个关键工艺流程步骤,简要

介绍 CFET 的一般结构,实际生产过程中的各工艺、膜层结构更加复杂。

本章小结

本章主要介绍了光刻工艺处于工艺流程中的位置、6T SRAM 的电路结构和工作原理以及三个技术节点中不同晶体管结构的简单工艺流程。晶体管的发展经历了平面结构、鳍形结构以及纳米板的 CFET 结构。本章以简要的流程图,大概介绍三种晶体管结构前、中、后段中一些关键层次的芯片制造工艺流程。每层光刻之前都会根据需要生长合适的薄膜层次,实际工艺流程中包括更复杂的薄膜生长、光刻、离子注入、刻蚀、炉管退火、化学机械平坦化等工艺。

本章对比了一种 14nm 技术节点中高密度 FinFET 和 3nm 技术节点中高密度 CFET 的 6T SRAM 基本版图结构,以说明 CFET 晶体管结构的特点。由于 SRAM 的原理是基本不变的,因此以 CFET SRAM 的基本单元截面图为例,同时搭配 HKMG 平面晶体管与 14nm FinFET 结构中部分层次的立体图,通过交叉学习,希望读者可以通过本章大概了解芯片制造的工艺流程、各光刻工艺层次的基本顺序和作用。另外,本章提到的多个技术节点中各种 SRAM 结构、工艺流程只是可能的示例,实际情况复杂多变。

参考文献



参考文献