

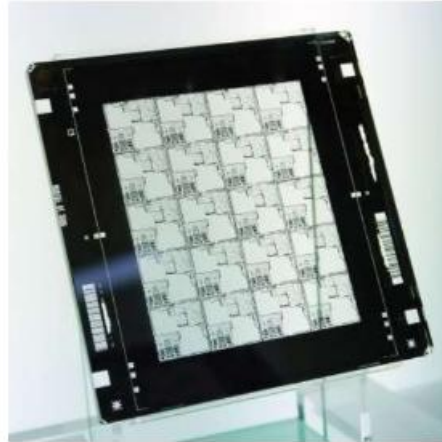
VLSI Design Flow: RTL to GDS
Dr. Sneh Saurabh
Department of Electronics and Communication Engineering
IIT-Delhi

Lecture 9
Overview of VLSI Design Flow: VI

Hello everybody, welcome to the course VLSI Design Flow: RTL to GDS. This is the eighth lecture. In this lecture, we will be continuing with the overview of VLSI Design Flow. In the earlier lectures, we have looked into the design steps that we carry out, from the idea up to the final layout. Once we get the layout, we need to send that layout to the foundry, and we finally get a fabricated chip. So in this lecture, we will be looking into the processes that are involved in taking the layout and finally creating or fabricating a chip out of it.

Now, this course is related to designing a chip or design. So why are we looking into the fabrication part? So, if we understand the fabrication part or the steps involved in taking the layout to the final chip, then we can appreciate the challenges involved in fabrication, and perhaps we can solve some of those challenges much more efficiently during the design step. Therefore, it is essential to understand a few basic concepts related to the fabrication of integrated circuits, and we will be covering those concepts in today's lecture. In the earlier lectures, we have looked into a fabrication task known as photolithography and we have also discussed that photolithography is a fundamental task that is used in fabricating an integrated circuit, and for photolithography, what we need is a mask.

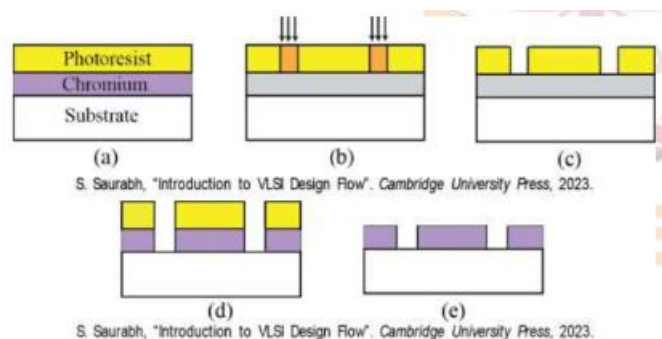
A mask contains the features of the design, and we replicate that feature on a substrate or on the chip. So, to start doing photolithography, we must have a mask. So, what is a mask? Just to recapitulate, a mask is the replica of the patterns on a given layer of the layout created on a substrate, which can be of glass or fused silica. Why do we create this mask? For transferring patterns during the photolithography step. So here is a diagram of or a figure of a mask. Now, we need to fabricate the mask before fabricating the corresponding integrated circuit.



Source:
https://commons.wikimedia.org/wiki/File:Semiconductor_photomask.jpg Peellden, CC BY-SA 3.0
 <<https://creativecommons.org/licenses/by-sa/3.0/>>, via Wikimedia Commons

Some of the typical steps involved in making a mask are data preparation, mask writing, chemical processing, and finally, ensuring the quality checks and adding protections. So, we will be looking at these steps in more detail in the subsequent slides. So, the layout that we create can contain complex polygon shapes. So, for mask writing, we want to translate the layout-specified information to a format that can be comprehended easily by the mask writing tool. So, what we do is that we convert complicated polygons to simpler rectangles and trapeziums, and this step is known as fracturing.

During data preparation for mask writing, we also augment the mask data to enhance the resolution. What are these resolution enhancement techniques? We will see this in subsequent slides. Now, how is mask writing done? So, in mask writing, we start with chromium and photoresist coated on a glass or quartz. So we start with a substrate, which can be of glass or quartz over which chromium is coated, and over this, there is a layer of photoresist. So, this complete structure is known as blank.



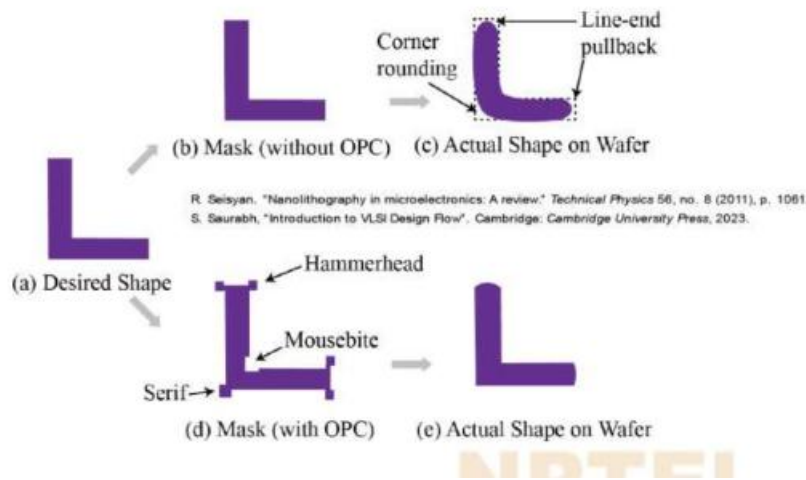
We start with a blank over which we write the required pattern, and these patterns are written either by exposing to the laser or electron beam. Now, when we write a pattern on the photoresist, what happens is that the properties of the photoresist change in the region which are exposed to laser or electron beam. As a result, when we subject this substrate with chromium and photoresist structure to a developer solution, the region of the photoresist, which was exposed to the light or a laser or electron beam, is removed on development, and we get the desired pattern on the photoresist. Then, later, what we do is that we etch out the chromium.

So, chromium will be etched out in the region from where the photoresist has been removed because chromium will be exposed in these regions, and we will get a pattern on the chromium also. Now, after we have got a pattern on the chromium layer then, we finally strip off the photoresist layer, and we get the required pattern on the chromium, and this mask now carries the features that we want. So, finally, what we do is that we do some quality checks after mask writing, meaning that we inspect for defects by scanning its surface and comparing it with a reference image. So, we will scan the surface of this mask that was created and compare it to the reference image to see if there are any defects. If there are any defects that are beyond a tolerance limit, then we repair them with the help of a laser.

Finally, we get the required defect-free pattern on the mask, and then we apply a protective cover called pellicle over the mask, and then this mask is ready for use for photolithography. Now, let us look at a few resolution enhancement techniques. Now, in photolithography, typically, we use a light of wavelength 193 nanometers. Now, when the feature size is smaller than the wavelength of light, diffraction effects and other non-ideal effects come into the picture and the image that we get on the wafer or the features that we get after photolithography are distorted. So, if we take a layout and make the feature on the mask exactly the same as the feature on the layout, and we carry out photolithography, we will get distorted features on the silicon.

Now, let us look at a few resolution enhancement techniques. So typically, in photolithography, we use light of wavelength 193 nanometers. Now, when we carry out photolithography using light of this wavelength and the feature size is much smaller than the wavelength of the light, then effects such as diffraction become very important. If we have a feature on our layout and we draw exactly the same feature on our mask because of diffraction and other effects, the feature that we will get on the silicon wafer or on the substrate will be distorted. So if we start with some feature and draw exactly the same feature on the mask, then the image that will be created and finally what we get out of lithography will be actually distorted. So what we do is that instead of just using exactly the same shape on the mask, we precompensate the mask with some features.

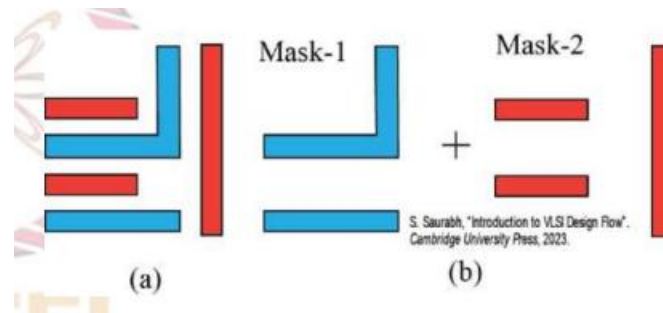
We actually add some controlled distortion to the mask as a result when distortion further happens to the features, what we get on the silicon is what we had desired. This technique is known as the resolution enhancement technique. So the mask is precompensated such that the features obtained on the mask are the same as desired, so we want this feature to be exactly the same as this feature, and to get that, we change the feature on the mask. An example of this technique is optical proximity correction or OPC and double- or multi-patterning. So, we will be looking into these two techniques in more detail.



So, when we print a pattern on a wafer, and the pattern size is smaller than the wavelength of the light, then it can undergo severe distortion, as illustrated in this figure. So if the desired shape is this L and we draw exactly an L shape on the mask, then because of the diffraction and other effects, the feature that will finally get on the substrate on the wafer will have distortions such as corner rounding or line-end pullback. Now, to avoid it, what we do is that we add features to the mask. So we add features such as hammerheads, serifs, or mouse bites to the features on the mask, and as a result, what happens is that when further distortion happens to these features, we get almost the same shape as the L shape that we wanted originally. So this improves the resolution of photolithography by compensating for errors that are introduced due to diffraction.

So this technique is known as optical proximity correction. The other resolution enhancement technique is double-patterning or multi-patterning. Now, due to the limited resolution of photolithography, printing closely spaced features is a challenge, meaning that if we have one feature and another feature is very close to it, then because of the limited resolution of photolithography, these two features can touch each other and these two lines can get shorted and the features can overlap each other. So, we can solve this problem by increasing the spacing between features printed at a time. So, if the features are very closely spaced, we do not print them together. In one step, we print this, and in the next step, we print this.

So, let us take a look at this technique in more detail. So, in double-patterning or multi-patterning, what we do is that we decompose a closely spaced layout into two or more layouts. So, this task is basically known as assigning colors to the features. So, to illustrate, suppose we had this layout. Now, in this layout, these two features are very close.



Then, we assign a red color to this feature and a blue color to this feature, meaning that we will be fabricating or carrying out photolithography for these two steps using separate masks or separate exposures. So, the first step is decomposing the features on the layout into separate colors and then fabricating features that are of the same color together. So, we use different masks and different exposures for layout features of different colors. Each exposure needs a lower resolution due to decreased feature or pattern density. So, in one mask, we have only the blue features.

So we see that the distance between the features has increased, and in another mask, we have these red features, and here also the distance has increased. So the good thing is that here we can fabricate very closely spaced features also using light of, say, 193 nanometers. The downside is that it will require a separate mask, and the fabrication step becomes more complicated. Now, once we have made the mask or we have incorporated resolution enhancement techniques in our mask, then we carry out wafer fabrication. So, wafer fabrication is the actual fabrication of a design on silicon, and this is carried out using a process based on photolithography.

So, it consists of hundreds of individual process steps. So this photolithography and other steps, for example, oxidation, diffusion, ion implantation, and many other steps, will be required in fabricating. There will typically be hundreds of individual process steps which will be carried out sequentially. So fabrication is done layer by layer. So we fabricate the circuit elements such as resistors, capacitors, diodes, and transistors, the active circuit elements, which are typically lying at the lower level of our integrated circuits.

We fabricate them using a set of processes, which are known as front-end-of-line processes. Then, over this, there are interconnections of metals. We fabricate these multiple layers of interconnects over the device layers using processes that are grouped

together and called back-end-of-line processes. So, we will be looking at these processes in more depth later in this course when we discuss physical design flow. So after we have carried out fabrication using, say, hundreds of individual tasks, we have got a die in which all the features have been incorporated. Then comes the step of testing, which was discussed in detail in the previous lecture.

Now, each die is tested and compared with the expected pattern. Now, if we find that some die is bad, then we discard them. We cut the dies from the wafer so that on a wafer, we can have many small dies. So we have individual dies, we slice out these dies from the wafer and if these dies are good then we proceed with the next step and the next step is the packaging step. So, in packaging, what we do is that we encapsulate these silicon dies in a supporting case, and once we have encapsulated these dies in a supporting case, then this is known as the chip. So, from the die, we get a chip after packaging.

So, what are the functions of the package? So we have this encapsulating case in which this die is placed, known as a package. Now, there are lots of functions of the package, and we need to consider them while designing a package for our chip. So, the foremost function of the package is to provide pins for connecting to the external environment. So, our chip will be actually connected with the external world, and these packages provide pins for them. Now, the characteristics of the package have a great impact on the delay of the signal entering and leaving the chip. So, the signal will be entering the chip and leaving the chip. Now, how much delay is encountered while entering a chip is very much dependent on the package pins, similarly on the output side, and therefore, we need to consider this important factor in our designing: how much delay will be there from the input pin of the package up to the pin where we are actually getting that signal and using it in our design. The package also allows dissipation of it, and this must be considered carefully.

So these packages have heat sinks, and if we do not use them properly, then our die will simply melt. It will not be able to withstand the heat that is generated during the operation of a chip. So, the thermal design of the package is a very important criterion, and the third is that this package prevents from mechanical damage and corrosion. So, the mechanical properties of the package are also very important. So, considering that packaging is basically performing many kinds of roles, there are various types of materials for packages have been designed or have been experimented with or are used in the industry, and there are various types of packages also. For example, there is a kind of dual in-line package in which the pins are on two sides of the package, and there is another kind of package known as the ball grid array type of package in which you have the pins arranged in the form of a grid on the one side of the package.

So these are just a few examples of packages. There are many other types of packages, some of which we will be looking into later on when we go into the chip planning stage.

So once the packaging of a die is done, then we do a kind of final testing. Why do we want to do a final testing? Because we want to ensure that the packaging step or packaging task did not introduce any error or there may be a situation in which the package is already faulty. So we want to check that after packaging also, our die is still functioning correctly, and the chip is giving the required results.

So this is what final testing is. Then, we can also do a burn-in testing, meaning that we subject the chip to a high voltage and high temperature. Why do we do this? We want that if there are any latent defects in our chip that were not discovered during the testing or manufacturing test step, those latent defects can become apparent when we do burning testing, and we want to avoid those chips from reaching our end user. Typically, if we look into the failure rate and the time, the failure rate is typically high initially, then it takes a kind of constant value, and then at the end of life again, the failure rate goes high. So this is kind of a bathtub kind of characteristic. So what we want to achieve in burning testing is that if there are latent defects and if there are issues of infant mortalities, those should be caught earlier during testing, and those products that will actually become defective at the end at the user side, are not reaching the end user at the first place.

That is the purpose of the burning test. At this stage, we also can do a binning. So, in the binning stage, what we do is that we classify chips that have been made from the frame design based on their performance. Now, when we design a chip, we expect that the expected behavior or performance after fabrication remains the same, but what happens is because of the process-induced variation and other things, there is a kind of spread in the performance of the fabricated chip. So what we want to do is that after fabrication, we do a kind of measurement on the chip using some on-chip delay measurement circuitry, and based on that, we assign different cost numbers or price points to different bins, meaning that if a chip has got higher performance, then it will be priced higher compared to the other chips or the chips of the other bins. So, after fabricating, we classify the chips into different bins based on their performance, and we utilize on-chip delay measurement circuitry to measure the performance, and we assign different price points to different bins.

That is why binning is done. So once we have done the binning and our chip is already fabricated, the chip is ready. We can either send the chips to the market directly or integrate them with other chips to make a system, and then that system can go to the market for the end user. So, this basically completes our discussion from the idea to the chip. So these are some of the important references which you can look into if you want to know more details. Now, to summarize, in the last six lectures, we have taken an overview of VLSI design flow.

We looked into system-level design. We looked into RTL to GDS design implementation, and then we looked into the verification task and testing. And in today's

lecture, we saw that after we have got the layout, how do we get the final chip? So, this completes the journey from the idea to the chip, and we have got a good overview of what the VLSI design flow looks like. Now, this completes the first part of this course. From the next lecture onwards, we will be going into logic design, and we will start with modeling the hardware using Verilog language.

So, this is the end of this lecture. Thank you very much.