VLSI Technology

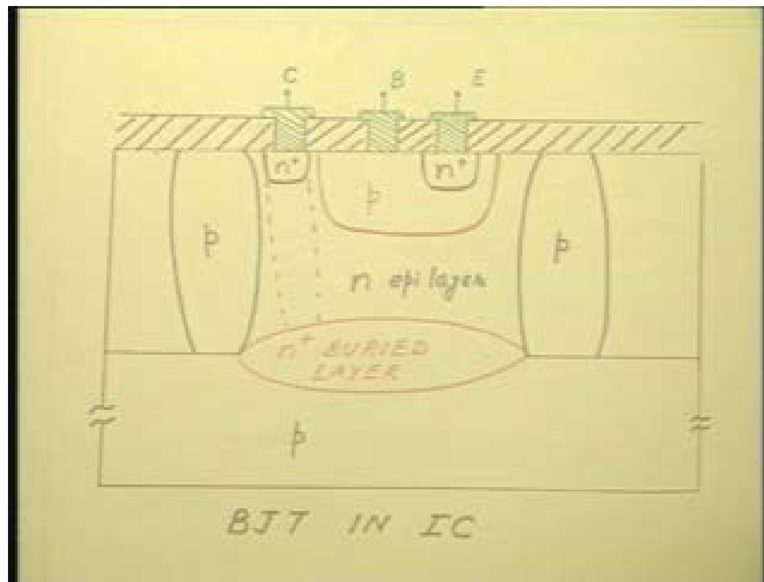
Dr. Nandita Dasgupta
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture - 32
IC BJT - From junction isolation to LOCOS

So, by now, we have completed all the unit process steps which are used in VLSI technology. So, now is the time for us to take a look at the two most important building blocks in integrated circuit technology namely, the bipolar junction transistor and the MOSFET and see what goes into making these devices. That is what particular processing steps with the details, what particular processing parameters go into making these devices and then see how the performance of these devices can be enhanced by suitably modifying the technology. So, you will see that with the advent of technology, all these devices have undergone a large metamorphosis, large changes, so that a bipolar junction transistor as it used to look 20 years ago, looks very different from the bipolar junction transistor today, because of the advance in the processing technology. Same thing is applicable for a MOSFET.

We will start with having a look at a bipolar junction transistor as it is in integrated circuit. We will first look at a comparatively older version of a bipolar junction transistor, point out what its drawbacks are and then see how by suitably modifying the process technology we can enhance its performance, right. So, here we go.

(Refer Slide Time: 2:58)



Let us take a look at this schematic diagram of a bipolar junction transistor here. This is something you are already familiar with. We started with a p-type substrate. Normally we used 1 1 1 substrate. Remember, it is a comparatively older bipolar junction transistor. So, in those days still 1 1 1 substrate was considerably cheaper than 1 0 0 substrate, because now you know that it is much easier to grow 1 1 1 crystal, right. So, 1 1 1 substrate used to be much cheaper that is why people started out with 1 1 1, about 10 ohms centimeter substrate and then you know, the first step was in forming this n plus buried layer. Of course, it has all the associated oxidation and photolithography to define the active area, but the goal of the first few steps was to realize the n plus buried layer and this n plus buried layer is needed to cut down the collector resistance.

This n plus buried layer can be formed by a diffusion process. It is n plus, therefore I need either phosphorus or arsenic or antimony, in order to realize this n plus buried layer. Now, the question is which one, which one of these dopant materials will be particularly suitable for this n plus buried layer formation? Very interestingly, it is not phosphorous, the most widely used n-type dopant material; it is not phosphorous that is used to realize this buried layer, because you see, this buried layer formation will be followed up by the

epi layer, by an epitaxy and the epi layer should have a much lower doping concentration than the buried layer.

So, the buried layer has essentially two requirements. The first requirement is that the surface concentration of this buried layer should be comparatively low. If the surface concentration is not low, then there will be considerable auto doping from this buried layer and the epitaxial layer doping concentration itself will get modified to a large extent. Problem will be particularly severe as we bring the device dimensions down. That is as the epi layer thickness is reduced, auto doping problem will be very severe and the other requirement for this buried layer is of course that it should have a low heat resistivity that is in order to bring down the collector resistance. So, it should have both - a low heat resistivity as well as a low surface concentration and phosphorous does not fit the bill, because it has a very high diffusion coefficient.

So, one is forced to use either arsenic or antimony. Arsenic and antimony have the advantage over phosphorous, in the sense that both of them have lower diffusion coefficient than phosphorous and in that case also antimony is preferred over arsenic, because you know arsenic is highly volatile. So, if you use arsenic to dope this n plus buried layer and then afterwards when you do the epitaxy, remember during epitaxy also the substrate has to be subjected to high temperature; so, in that case if you used arsenic there will be large amount of out diffusion from this buried layer. You do not want that that is why antimony is the preferred choice for this n plus buried layer formation. So, this n plus buried layer is used, in order to lower the collector resistance.

After the buried layer formation of course, you go to the n epitaxial layer and then in order to isolate this particular transistor from all adjacent devices, we must have a deep p diffusion going all the way down to the p-type substrate, so that this particular transistor is isolated from the adjacent regions, where other transistors are housed. This is, as I said, it is a comparatively older bipolar junction transistor which uses junction isolation, junction isolation. That is this reverse biased pn junction is used as the method of isolation.

Now, let us look at this junction isolation technology, because immediately after buried layer, we must concern ourselves with the junction isolation; with the isolation and the junction isolation. Now, you see, I am still taking about the older day technology. So, these junction isolations were achieved again by diffusion process. So, you have to carry out a deep diffusion all the way down to the p-type substrate. Now, in doing so, you see, the deeper the junction is, there will be more lateral spread. You know that lateral diffusion is about 80% of the vertical diffusion. So, since your p-type diffusion has a considerable depth that means it extends in both ways as well, so even if your original isolation window was small, there will be considerable spread on both sides and therefore, you will have larger area requirement. That is one drawback.

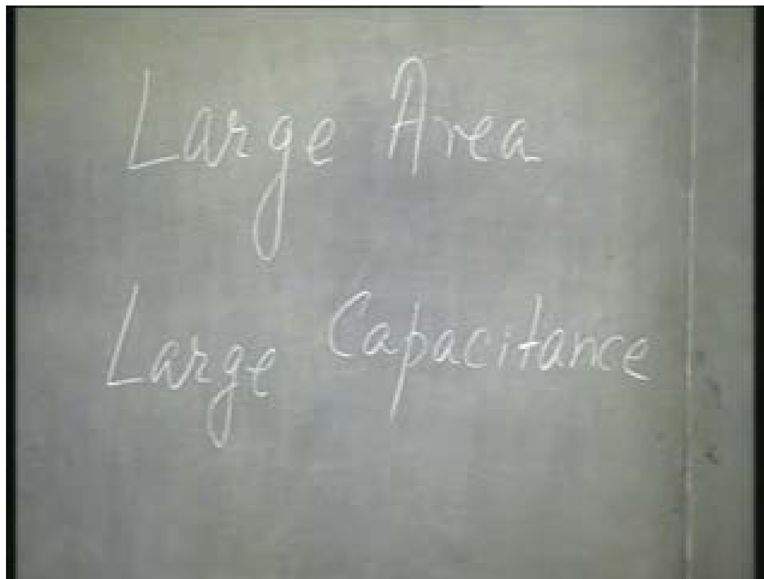
The other drawback is of course, you see, this junction isolation is p-type and my base is also p-type and in between I have the collector n region. So, along with my vertical npn transistor, I have a parasitic pnp transistor comprised of the isolation, the collector and the base of my active transistor; pnp, a lateral parasitic pnp transistor. Obviously, you want this lateral parasitic transistor to be least effective; you do not want it to be effective. How can you suppress it? By suppressing its gain. How do you suppress its gain? By increasing the base width.

This is a lateral transistor, pnp. So, this is its base, right. That means the base of your actual transistor should be a considerable distance away from the isolation p-type, in order to keep the base width of the parasitic pnp transistor large. So, the practice distance used to be 20 microns. If you have a transistor with a 20 micron base, it has absolutely no gain to speak of. So, you do not have to worry about the parasitic effects of this transistor. But, you pay a very heavy price in terms of real estate, right. Just as in real life, in VLSI technology also, real estate is the highest priced commodity. The silicon area is limited. I can have, may be a 6 inch diameter wafer today or may be 8 inch diameter wafer today. But still, my requirement is large. I want to put a large number of transistors in the given silicon area. So, I cannot afford to waste 20 micron here, just so that my parasitic transistor will not disturb the circuit performance. So, this is one major drawback of the

junction isolation technology, namely the large area requirement. This is so far as area is concerned.

On the other hand, as we move towards faster and faster circuits, we need to have more and more high speed transistors, then the limiting factor on the speed of your circuit comes about from all the parasitic capacitances, because finally it is the RC delay which governs the speed of the circuit. So, this, all parasitic capacitances will have adverse effect on the speed of the transistor and if you have this junction isolation, then you have a large junction capacitance associated with the collector, n and isolation peak. The pn junction capacitance will tell adversely on the speed of the transistor. So, you see, basically the pn junction isolation technique has two major drawbacks.

(Refer Slide Time: 13:39)

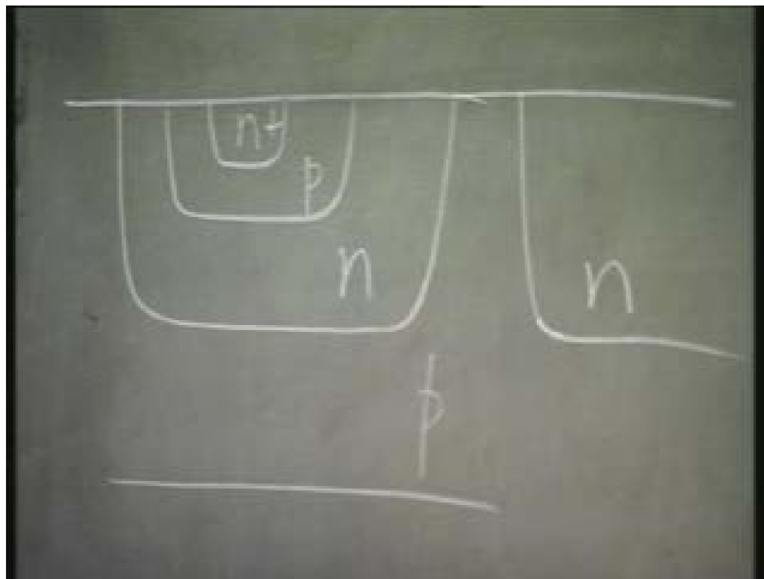


The first drawback is the large area requirement and secondly, it has the adverse effect on the speed by having a large capacitance. So, as the device technology matured, as our requirements became more stringent, as the circuit complexity increased that is one have to put more number of transistors in a given area and as the speed requirement also increased, it was found that this pn junction isolation technique may not answer to all our requirements. So, then, there were a few modifications.

I will first talk about a few modifications in the junction isolation technique itself, but they are more of historic importance than anything else, because nobody uses those techniques in those days, in these days. But still, let us see how in junction isolation they tried to modify something, in order to achieve either a lesser area requirement or in order to improve the speed and at what price? Then we will see, why these techniques did not answer to both our requirements and what needed to be done afterwards.

So, one modification that was thought of was instead of having the epi layer, n epi layer, people tried to have three diffusions in a p-type substrate. You do not have an n epi layer, instead of that you have a totally p-type substrate. In that you try to have three diffusions, something like this.

(Refer Slide Time: 15:48)

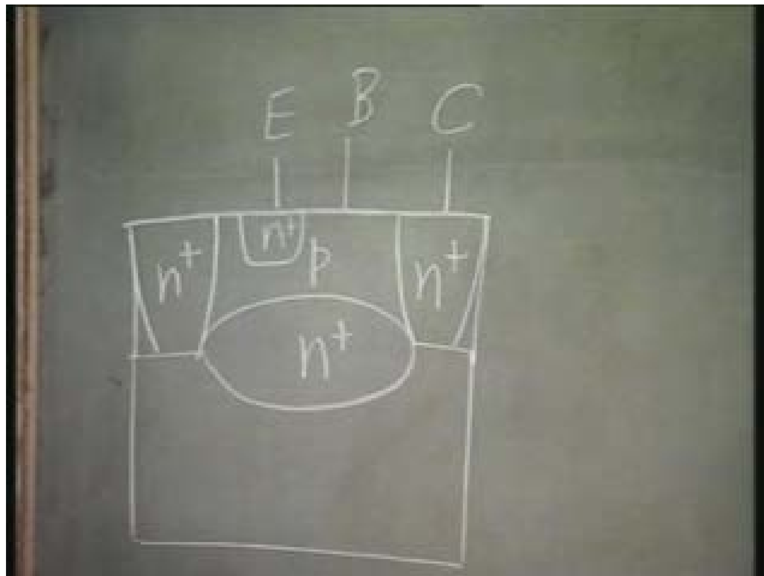


You start off with this p-type substrate and then you have your first n diffusion; this is going to act as your collector. Then, you have a p diffusion, this is going to act as your base and then you have … this is going to act as your emitter. So, I have emitter, base, capacitor all of them are housed in the p-type substrate and I have the next adjacent transistor housed in another n diffusion. So, you see, they are automatically isolated. I do not have to have this large p isolation diffusion and because this p isolation diffusion in

this particular case had to be very deep and therefore, it resulted in a considerable lateral spread, you could save some area. So, some improvement in packing density was achieved if you have these three diffusions; some improvement in packing density, because you no longer have to have a p isolation diffusion which will result in considerable lateral spread. But, at what price? The price you pay is in terms of collector resistance, because now you have three successive diffusions. Obviously, the n doping concentration has to be the smallest; of these three regions, the n doping concentration has to be the smallest, because you are doing successive diffusions, right; so, your collector resistance will be considerable high. So, even though you get some improvement in packing density, your collector resistance of the transistor will be extremely high. So, this was not found to be very suitable, even though you can get some improvement in packing density.

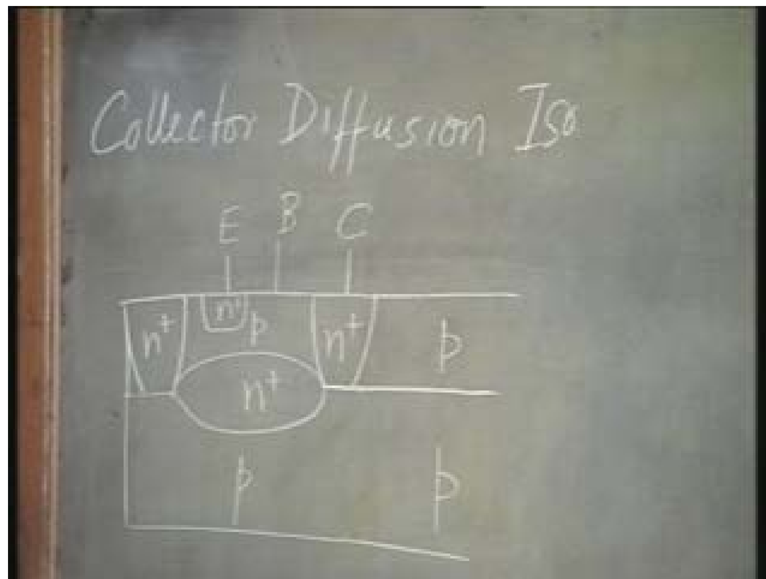
The other technique that was used again in junction isolation, the other choice was called collector diffusion isolation; CDI, collector diffusion isolation.

(Refer Slide Time: 18:41)



In this what you do is you start off as before with the p-type substrate and then, you have the n plus buried layer. Up to this, it is identical to this older bipolar junction transistor technology. The difference is after this. Now, instead of having an n epi layer, you have a p epitaxial layer and then, you diffuse n plus. Instead of diffusing p plus, you diffuse n plus and then, you can have your … So, this is now your emitter, this is your base and this is the collector.

(Refer Slide Time: 19:58)



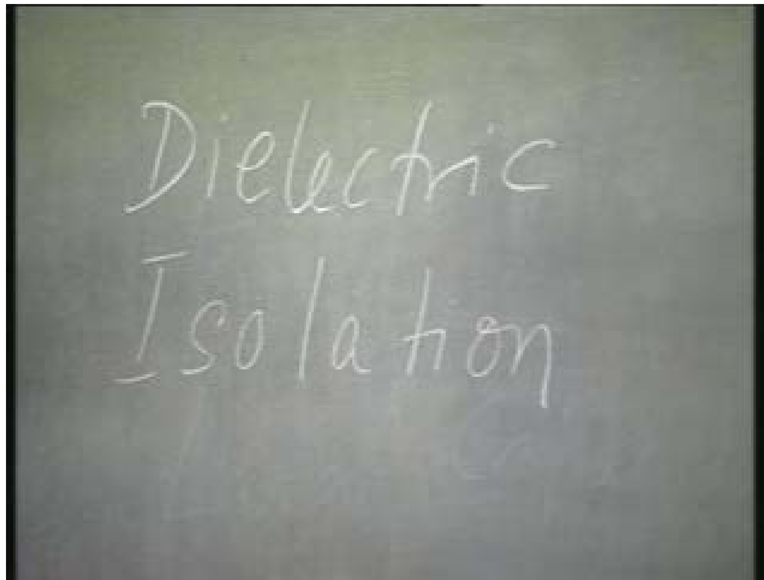
So, if I look at the next device, this is p-type substrate, this is p-type substrate and this is p-type epi layer. If I look at the next device, next transistor here, that is still isolated by the reverse biased np junction. That is each transistor is surrounded by n plus region, by this n plus p. Now, collector is n plus. I have throughout n plus collector and this collector substrate junction is isolating the two adjacent devices. The reverse biased collector substrate junction is creating the isolation. So, you see this is called collector diffusion isolation; collector isolation diffusion or CDI in which case you started out in a very similar way, started with the n plus buried layer, but then instead of having an n epitaxial layer, you have a p epitaxial layer and instead of having the isolation p plus diffusion, you have n plus collector diffusions.

So, this n plus collector diffusions together with the n plus buried layer is serving the dual purpose. It is performing as a collector and also doubling as an isolation region and here of course, you have overcome the first problem which we had in the previous three diffusion technique that is the problem of collector resistance is sorted out. Now, your collector is n plus, therefore the collector resistance is very small; but, even that can create a problem. Since collector resistance is very small, you may have punch through in the base and the base collector is reversed bias. Because the collector is very highly doped, entire base can get depleted. So, punch through in the base may occur.

So, what is the remedy? You have to increase the epi layer thickness, if you want to avoid the punch through. So, you see, even if we try to modify the junction isolation technique, the basic problem still remain there. So, junction isolation is still used; it is still quite a viable technology particularly for analog devices, analog transistors; junction isolation is used. But, people began to realize more and more that there must be a better alternative to junction isolation. After all, what do we want to do? We want to isolate the two transistors. That means we do not want any electrical connection between the two transistors. In this case, we are doing it by a reversed bias pn junction. For junction isolation we are doing it with a reverse biased pn junction, because you know in a reverse biased pn junction, current does not flow unless you have exceeded the break down voltage.

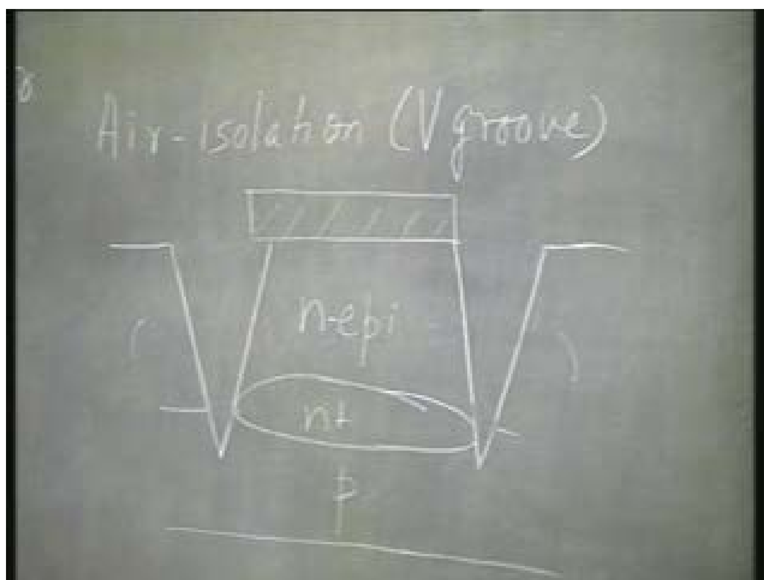
Isn't there a better technique? What about putting an electrical insulator between the two transistors? If I can have an electrical insulator in between two transistors, no current can flow. So, what are the insulators that come to your mind? It could be air, it could be oxide. So, another interesting technique is to use a dielectric material for isolation. Instead of using a pn junction, you go to a dielectric material for isolation. So, from junction isolation, we have moved on to dielectric isolation.

(Refer Slide Time: 24:55)



Let me again remind you some properties of silicon crystal. Remember that we can get anisotropic etches which will etch preferentially. That is most anisotropic etches, they do not etch 1 1 1 direction, they will etch all other planes, so that if you have a 1 0 0 substrate and you subject it to an anisotropic etch, you will have V grooves etched in it, remember. Let us see whether this can be used.

(Refer Slide Time: 25:46)



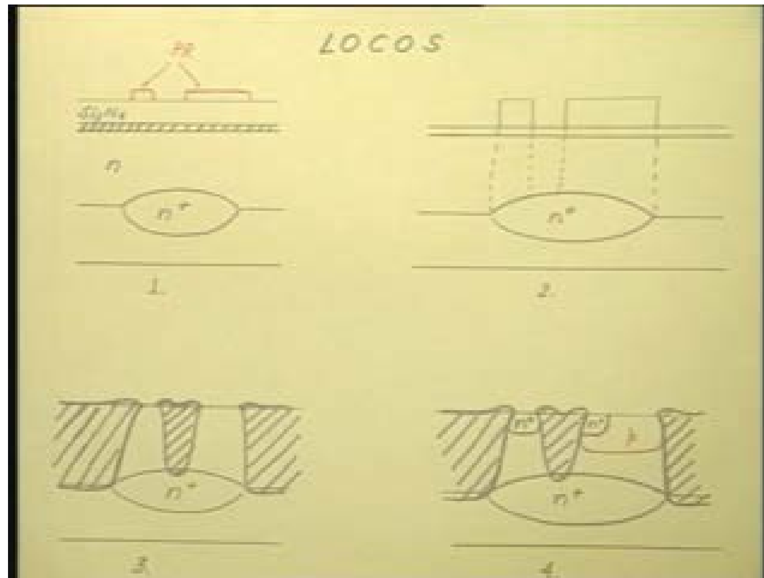
Let us say, I started off as before with the p-type substrate and I have this n plus buried layer, right and now, next I have grown the n-epitaxial layer on top of this, right. Let us assume that the substrate was 1 0 0. The epi layer is also grown in the 1 0 0 orientation. Now, let me protect my active region like this, using oxide or photoresist or whatever; by a suitable protection, let me protect my active region and now, let us subject it to an anisotropic etch. So, what will happen? I have etched V grooves in the substrate. So, you see this is my first transistor. This is the area for my second transistor, this is the area for my third transistor and all of them are isolated from each other by this air V groove. There is no path for electrical conduction, right. This is n plus p, reverse biased n plus p and there is no path for any current conduction. So, this is air-isolation, also called a V groove technology.

Are you all completely happy with this? What are the drawbacks of this technology? Well, the first drawback is now you are limited in your choice of the substrate. Remember, previously we started off by using a 1 1 1 substrate. Now, if I have to use this air isolation technology, I cannot use 1 1 1 substrate; I am restricted to using only 1 0 0 substrate, because the V groove can only be etched in 1 0 0 substrate. The other problem of course, is the problem of topography. That is it is no longer a planar technology. Look at the surface; it is all ups and downs and with this V grooves etched in it, if you try to do any photolithography you will have a very tough time indeed. So, the major drawbacks of this air isolation is, one - restriction on your choice of substrate; you have to use only 1 0 0 substrate and the other point is loss of planarity. The surface topography is non-planar. So, even though dielectric isolation seems a very interesting alternative to junction isolation, we saw that our first choice of the dielectric material, air that is, it does not answer very well to all our requirements.

So, the next dielectric what comes to your mind is of course, silicon dioxide and yes, it is the silicon dioxide isolation that is the most prevalent technique today. It is used not just for bipolar junction transistors, but also for MOSFET. In fact, it was first used for MOSFETs and then from MOSFET technology, BJT borrowed it and used it. So, now it

is a selective oxide isolation or referred to as LOCOS, local oxide isolation. That is the most prevalent technique today.

(Refer Slide Time: 30:29)



So, if you look at the next card, we will find how a LOCOS isolation is taking place; LOCOS that is local oxidation. The most key factor, the most important, the key factor in this is the term local oxidation. We want some region of the semiconductor to get oxidized and we want some other region, the active region for example, to be prevented from oxidation. We do not want oxidation to proceed in the active region. We want oxidation to proceed in only the region surrounding the active transistor and let me recall to you a very interesting property of silicon nitride which makes it a very interesting material in VLSI technology. Silicon nitride does not allow oxidation. Silicon nitride is resistant to oxidation; a region protected by silicon nitride cannot get oxidized. So, this is the key in the LOCOS technique.

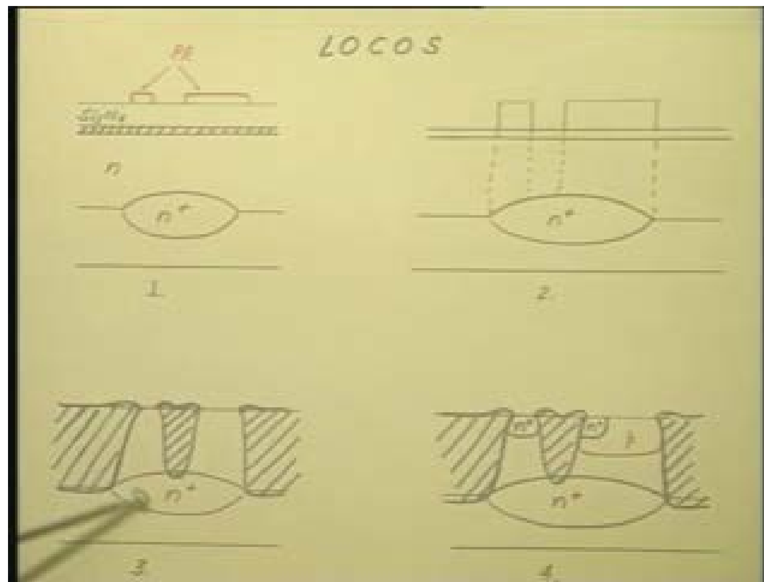
Let us take a quick look at the LOCOS technique. Let us look at figure1. What do I have? I have as before the p-type substrate, I have the n plus buried layer and then, I have the n epitaxial layer. So far, nothing has changed; so far, my technology is exactly the same as the older day bipolar junction transistor. After this of course, what have I done? I have a

thin layer; this black hatched region, this is the thin layer of silicon dioxide which is called the pad oxide. On top of the pad oxide, I have silicon nitride Si 3 N 4. This has to be chemical vapour deposited silicon nitride, stoichiometric silicon nitride. The plasma deposited ones do not work very well. So, most people use chemical vapour deposited silicon nitride. So, on top of this you have chemical vapour deposited silicon nitride and then, you have a photoresist pattern on that. That is some regions of this silicon nitride are protected by photoresist. So, this is your first figure.

In the next figure, using this photoresist masks I have etched silicon nitride from the remaining portions, so that I have silicon nitride only existing in these two regions and underneath there is this pad oxide as silicon nitride is etched from all other regions. So, now you see, the equation is very simple. The regions which are not protected by silicon nitride will get oxidized. If I subject it to oxidation, the regions which are not protected by silicon nitride will get oxidized. However, there is a slight problem. You remember, when silicon is subjected to oxidation, it consumes silicon. But, because of the difference in the density and the volume, if I have 1000 Angstrom of silicon grown, it consumes only about 450 Angstroms of silicon. That means 550 Angstrom is on top of the erstwhile surface, while 450 Angstrom is below.

If I simply use this silicon nitride as the mask and subject it to oxidation, then the regions where oxidation proceeds I will have silicon dioxide coming on top of the surface. Again, there will be loss of planarity. Do you understand the problem? So, instead of that, what is done is one additional step. As I have shown by this dotted line, you etch a little bit of silicon. That is let us say, if you want to, whatever is the oxide thickness you want, you etch it down by about 0.55 times of that. Reactive ion etching is usually adopted, so that you have highly directional etching. So, this is called recessed oxide, recessed oxide. That is you first create a recess in silicon and then you oxidize it. So, you etch it down to certain extent and then you subject it to oxidization, so that your calculation is such that the top of the oxide comes only up to the surface. Etching is done to realize only that much of recess, so that the top of the oxide comes only up to the surface.

(Refer Slide Time: 35:42)



So, after this when you carry out the oxidation, this is how the surface looks. You carry out the oxidation and then you remove silicon nitride from these remaining portions also. So, you see, this is your n plus buried layer and on top of that you have realized these two regions. Now, in one of this, now you carry out your base diffusion. Note a very interesting thing. The base now can be stacked against the two oxides on both sides. There is no question of having the base separated by 20 micron or whatever distance, as in the junction isolation case. The base now can occupy the entire region in between the two oxides which means a tremendous improvement in the packing density. Same thing is for emitter also. You see, emitter is also stacked against the oxide here and you use this region in order to realize the collector contact. This n plus region is going to be used as the collector contact, this n plus as the emitter, p is the base and then, underneath you have the collector and the buried layer. So, this is your collector path, path of the current flow is like this. You are forcing the current to go through the buried layer, so that the collector resistance is cut down.

So, this is a recessed oxide LOCOS isolation technique. Recessed oxide is used in order to improve the planarity of the device. Of course, without recessing also you can use the oxide isolation technique. In that case, the surface, the top surface of the oxide will rise

above the actual device surface. So, there will be a loss of planarity. But of course, if you do this recessing, then you have a slight non-planarity on the surface which is caused because of lateral oxidation. See, when you are etching it down, two surfaces are exposed. One is the horizontal surface and the other is the side wall.

So, oxidation progresses both on these surfaces. Therefore, there is a slight encroachment under the mask which is actually called the bird's beak problem, because it looks somewhat like a bird's beak. But a bird's peak problem is not very severe as far as the bipolar junction technology is concerned. We will discuss in more detail about bird's beak problem when we discuss MOSFET technologies, because there it will be more severe. We will discuss there about bird's beak problem and how those problems can be minimized. For bipolar junction transistor, however essentially this LOCOS technique is used for isolation that is protect the active region by using silicon nitride, etch by a calculated amount, so that the resulting oxide comes only up to the top surface; carry out the oxidation in order to realize the isolation regions and then have your base, emitter and collector contact diffusion.

The advantage is tremendous improvement in packing density without having to sacrifice the collector resistance at all, because even now you have a very low collector resistance; you are forcing the current in fact to go through this buried layer. It cannot go like this, because these two regions are isolated again by this oxide. So, only path for the current is to go through this n plus region. So, you have achieved both your objectives - a low collector resistance as well as a very high packing density. Of course, the oxide isolation using LOCOS is a very popular technology today, but even further improvement has been achieved and these are done by more complex processes. When the device dimensions really become very small, sometimes even LOCOS cannot answer to the description, because you see, when the device dimensions become very small, then this encroachment problem becomes a bit severe, the birds beak problem that is how the oxide is encroaching underneath the regions protected by silicon nitride. So, then, even that may become a problem. So, in that case, one uses a newer technology, which is called a trench isolation. We will discuss about the trench isolation in the next class.