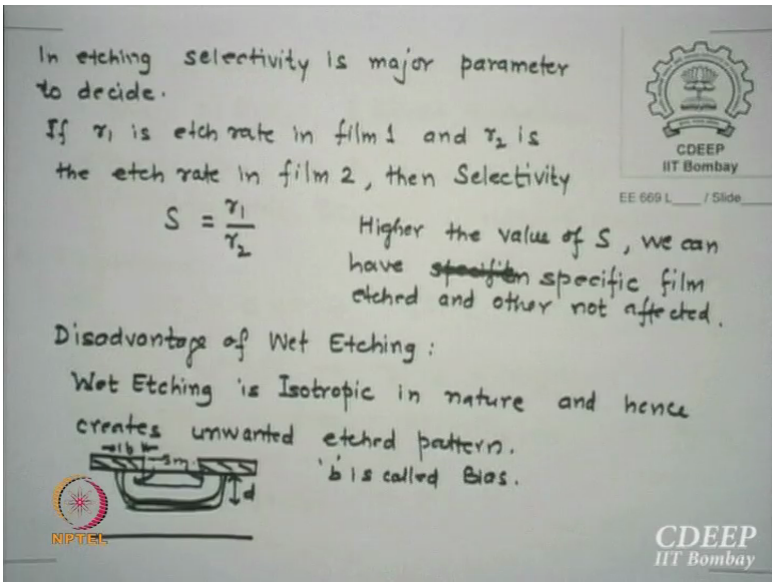
Fabrication of Silicon VLSI Circuits Using the MOS Technology
Professor A. N. Chandorkar
Department of Electrical Engineering
Indian Institute of Technology Bombay
Lecture 28
Etching in VLSI Processing & Back End Technology

Before we start with the last chapter which is back end, let me complete what we were discussing last time about etching in VLSI and we discussed last time about the wet etching and today we will look into dry etching quickly and then go to back end. One of the problem which I said last time that if you are doing a wet etching most wet etchants are isotropic in nature and when they etch they not only etch vertically down but also in the lateral side.

And that gives you whatever is your mask pattern the etched area is more than what the mask actually was asking. Now the additional on both side areas which is etched has been given in b and it is called bias. And if we look at this, this we have done last time so I am just trying to be quickly on what we did.

(Refer Slide Time: 01:12)

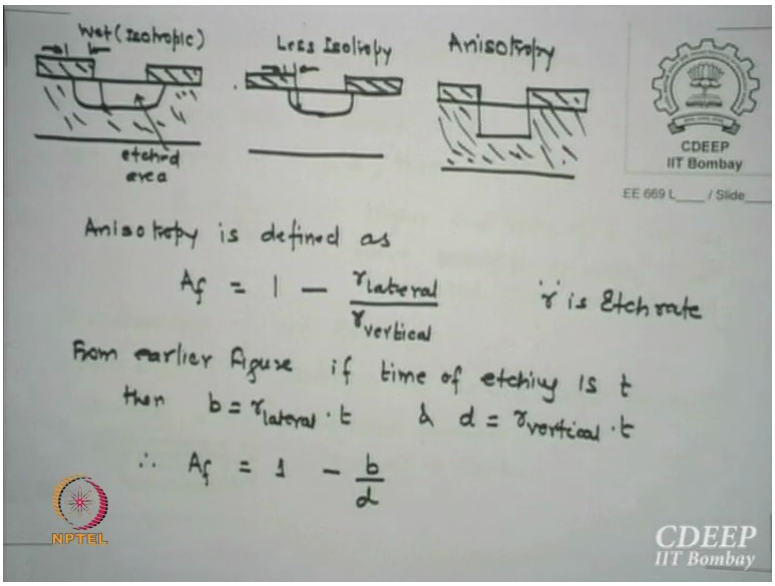


We did say that if this is larger b are much larger so that the mask distance, the width is much the total film which is etched is much larger than the mask width and this is essentially a bad area. I may do a little better etching in which less amount of lateral etching is done so I will have a better compared to this. How ideally I want anisotropic etch in which whatever is the mask width same as the etched area. Now we defined an anisotropic term as A f which is 1 minus etch rate in lateral direction upon etch rate in vertical direction.

And if you see if etch rate in lateral is zero that means nothing goes on the left side or right side then A f is 1. So the ideal profile if you want then A f should be 1. But if r lateral is positive and sufficiently close to vertical it can be point 1 to point 1 is the value of A f can vary. Smaller the A f, worst is the dimension you are getting and larger the A f, better dimensions are what exactly mask wants we are transferring.

Now this proportional if you note etch rate is into time is equal to the distance b on the lateral side and on the d which is the depth of the film which is r vertical into t. So, essentially what I am saying that A f is 1 minus b by d. So this is the film thickness, this is the bias and A f decides how much is the value you are getting. D is the film thickness vertically down. B is the bias lateral. So b by d is essentially 1 minus b by d is the anisotropy coefficient and we are expecting that to be as close to 1 as possible. There is little maths I did.

(Refer Slide Time: 03:26)



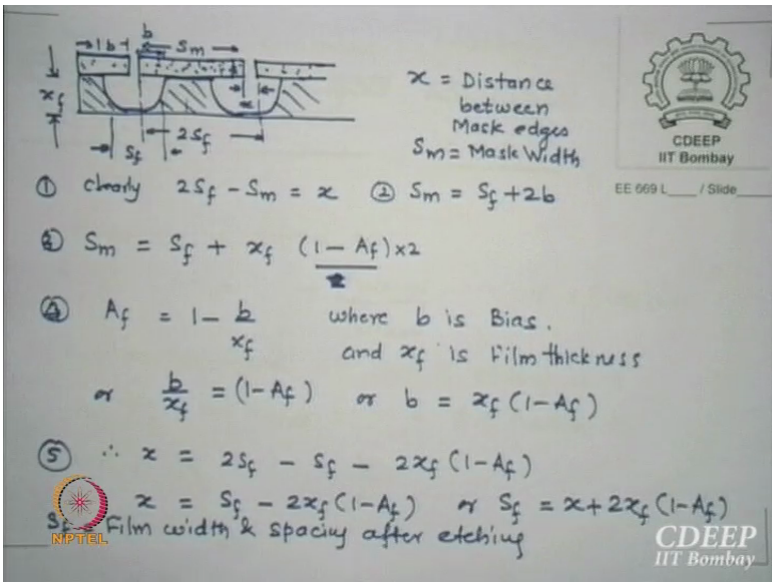
Of course this is given in (())(03:26) so you may even see later. Here is something what the pattern look like. This S m is the mask size you want to etch below that, okay. This is another mask area and in between is the distance x which is going to be etched, okay. Between the mask edges the distance is x which is where the etching is going to be done, okay. So we now define some terms. If this is my etch profiles on both sides of x, I define the term S f as the film width which is taken from of edge this mask to the edge of the other mask.

It is called pitch of that. This is the total if you look at it from this point to this point is 2 S f, okay. S f is the film thickness shown here. S f is whatever is this much part. So from here to here it is 2 S f. This is your S m, this is your b, this is your b. So if I do little maths on these

one can say 2 S f minus S m is the edge to edge mask distance x which is what we are etching actually, x is what we are etching. We also see is S m is S f plus 2 b, b b so S m plus 2 b please remember S m minus 2 b is the S f.

What is S f is the film width which you actually etch edge to edge. So, S m is S f plus 2 b. We do little calculation again as we did for A f we will come back to this later. A f is equal to 1 minus b by x film thickness. B we wrote but here it is I said it is x f so b upon x f 1 minus of that is A f. So b by x f is 1 minus A f so b is x f into 1 minus A f. This is simple maths. I will come back to it.

(Refer Slide Time: 05:29)

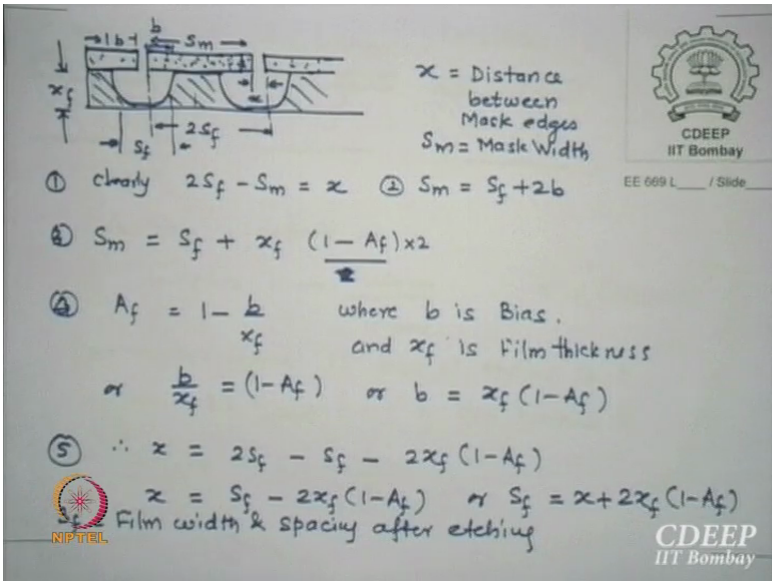


S f is the distance shown here, okay. So you calculate that will be 2 bases. So if I derive b is equal to x f 1 minus A f, I substitute that in this equation S f plus 2 b so I write S m is equal to S f plus x f 1 minus A f into 2. This equation has been now reduced to this or rather expanded to this and then from here I can get a value of x which is 2 S f minus S m is x, 2 S f minus S m. S m is this expression so if I write this I get expression x f, x is equal to S f minus 2 x f 1 minus A f.

What is that we are trying to say from this? What is the ideal etching we are looking for? S f should be same as x. So if A f is equal to 1, x is exactly what edge to edge you want same etching it will go. If it is not so then whatever you wanted actually it would be larger than x and something additional window will be opening, okay. Now this film thickness which so one can see from here if I write this S f is x plus 2 x f this formula, I can make S f equal to x in what conditions? One is A f equal to 1.

What is other possibility? If the x f is zero or smaller or thinner, smaller the x f but that is obvious if the film thickness is very low the lateral etching will not have enough time to really go deeper. So essentially we are saying if the film is thin then it is not that great to get x equal to S f. But if the film is thicker then the lateral etching may happen and there x will not be same as S f.

(Refer Slide Time: 07:21)



So possibilities that x is equal to S f is very unlikely unless I make A f equal to 1 which in most cases would be very difficult but we will try. I may get A f is up to point 8 to point 9 and can get but never we can get 1. Some lateral etching will happen, okay. Why it can happen? Even if something is going down at the edges it will hit something called at the edge there is a field they say so ions will be pushed even inside, a bit of them. So they will be additional b always will be there which will not be zero so A f can never become.

So the figure which we draw last time, this is only ideal figure that A f is equal to 1 which is very unlikely to attain. Now why we are worried about because the people who designed the mask they have two parameters S m and x. Edge to edge what is the distance between the mask and the mask size itself. They are fixed but now we are going to etch it and you say okay the actual film thickness which you etched actually giving larger than what x wanted.

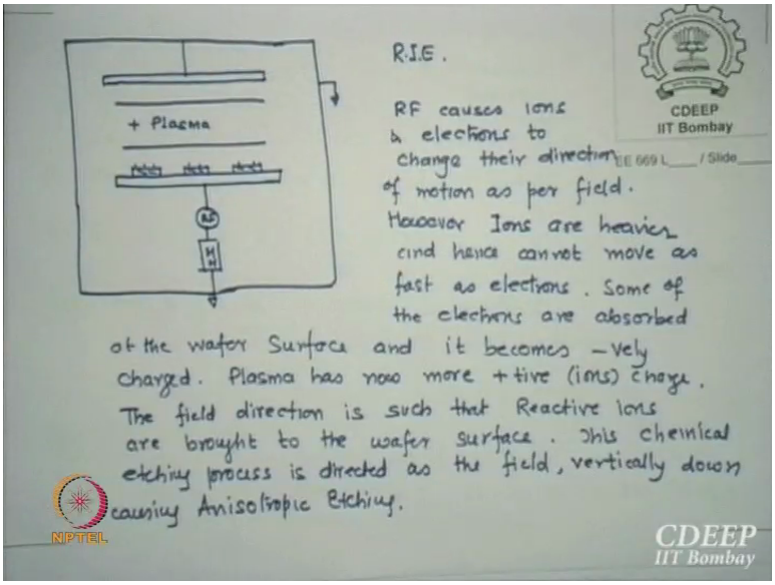
That means you have air in there because there is a next device may come closer to it. It may happen both devices may merge in some cases because etching may overtake the other side also, okay. So such a situation we want to avoid. So what you are telling the designer. Keep at least away this much. This is called design rule, okay. We are actually telling them never get

1 so go at least this distance. But at least larger the distance I ask them to go what does that mean? The area of the chip will start increasing for the same function.

So designer, his effort is to see making it compact. As much transistors and interconnect I can push per unit area that is my aim. And now you are going to tell I have a problem. So the technology must try to achieve as good A f as is possible and one possible which is done is using what is called as reactive ion etching. So what is our ultimate aim? I repeat x equal to S f is what we are looking for, okay. Film whatever width etched same should be the edge to edge mask which is window for us.

So what is the reactive ion nature? A reactive ion nature, this formulation is given in books so there is no great thing I am talking about. Can you see this figure slightly different from the plasma deposition system I made? Yes, in this case the wafers are connected to R F source and the top plate is essentially grounded. In other cases what was there? The source was here and whatever is the target was kept here, okay. Is that clear? The wafers are now target and the upper is grounded plate, okay.

(Refer Slide Time: 10:38)



Now, one can see from here the way it happens if I apply R F across anode and cathode, the gas may be discharged because of the field is large enough. However if you see ions in electrons, electrons are more mobile compared to ions. Ions are heavier mass, the velocity will be smaller. Electrons are lighter, they will move faster. So it may happen that electrons may first hit the cathode or the substrates earlier than ions can come and they may actually charge the substrate by negative charge. Is that correct?

That is why that cathode potential higher because cathode may become much more negatively charged. Is that clear to you? But since some electrons are left this from the plasma but there is still not all electrons are left there. So this is still a plasma but now it becomes little more positive because negative electrons have left. Therefore if you see other day figure there is a plasma potential positive shown. Is that clear? Now it is clear.

Faster electrons move out some of them by the time they absorb and secondary the plasma gets into, by the time the R F frequency allows it to go opposite polarity. So this plasma getting positive charge with reference to cathode is very dominant in fact which allows now whatever is the plasma ions they will be accelerated by the field across this and will hit the substrate. Is that correct? It will hit the substrate. The resist normally stops many of these ions like ion implantation we see resist actually can stop most of the ions.

In this case also if you are resist there during etching, ions cannot remove resist but wherever is the window develop ions can get inside. Is that clear? And they will hit whatever film there may be a (())(12:42) nitride or metal or anything which is there it will start attacking. Now there are two possibilities exist in this if the gas which are use is not just argon or some kind of a neutral gas, okay. If let us say I use a gas which is Freon, okay. C F 4 or whatever required to and it may release fluorine ions.

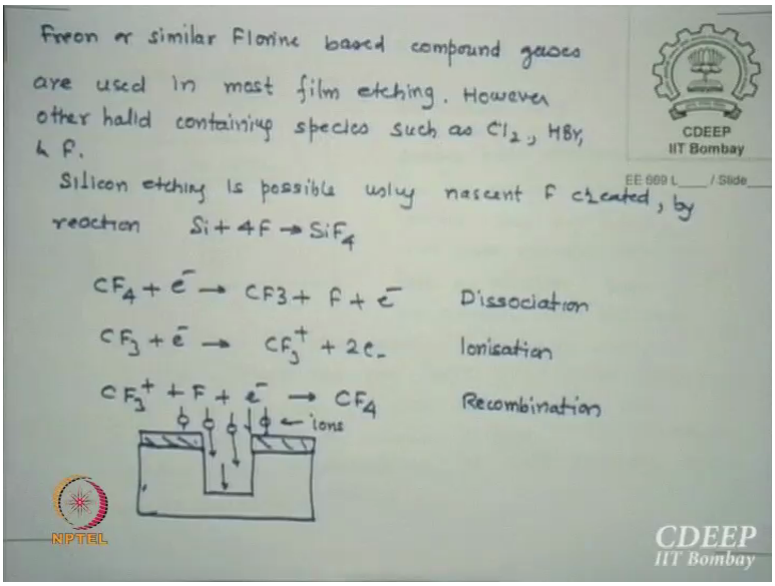
But fluorine at least for silicon diodes, silicon nitride and some metals as well fluid attach very heavily. What is this reaction will be? Chemical reaction. Fluorine is actually chemically reacting the film down but at the same time the fluorine ions are moving in which direction? Vertically downward because of what? Because of the electric field sets in there, okay. So ions are accelerated but they are active, they are charged and also they are reactive to the film itself. So this is what kind of etching we are doing? Chemical as well as ionic. Is that correct?

What the two kind of etching I did? Chemical etching because fluorine is attacking as well as the bombardment is done which is causing energy to hit the target and remove there from the like a (())(14:03). So what two process is going on in the case of reactive ion etching. The reaction is taking place from the chemicals and bombardment is taking because of ion being energetic vertically going down. Is that correct? If I only use argon then what will be it called? (())(14:20) because then there is no reaction between argon and this.

So that is the difference between (())(14:27) and R I E. Only difference is that I say the gas used is reactive kind. If you look without going into too detail of chemistry look for yourself. A Freon gas when it picks up some electron it becomes C F 3 plus fluorine plus electron.

Then C F 3 plus electron is this. C F 3 plus F e minus C F 4 this is called dissociation, ionization, recombination and because of this fluorine ions actually travel down, react with the film below and also since they are vertically down they also hit. Ion energy is given so they also do sputtering. So there are two mechanisms going on in etching. What is the first? Flow reaction and second bombardment of fluorine ions, okay.

(Refer Slide Time: 15:15)



These two mechanisms lead to additional etch rate. Please look at it. There right two etch rates now going on, one essentially because of the reaction, the other because of ionic bombardment. Is that clear to you? So normal R I E has two kinds of etching mechanisms which we must model in case we want to find the actual etch rate. However many times I am not very keen to have bombardment. I am not really looking for anisotropy.

Where do you think I do not need anisotropy during etching? After I do lithography I developed the pattern then I do R I E etching, okay. But what is now left after etching? The resist everywhere but that I am not etching selective. I want to remove photo resist everywhere. Is that correct? I am trying to remove photo resist everywhere. This is called ashing, okay. This is called ashing. So I want to remove the photo resist. For doing this I do not really need ions for that.

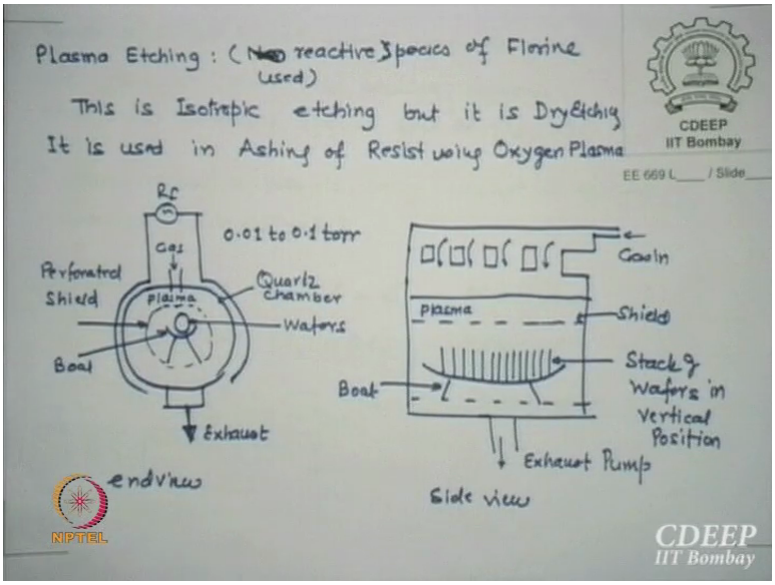
I do not need directional ions, I only need fluorine to be available or something else to be available which can etch photo resist. so oxygen plasma is good enough. Without any direction it can remove the resist. This is called ashing and you have written down these reactions. OF course these are also given. I do not know about in (())(16:57) but in every etching book you can see this same reactions. Which gas is now banned ? Freon is banned gas now, okay.

So which gas is used? In air conditioning which gas is used? Even chlorine is now getting banned. So now the gas probably is going to use is S 2 F 6 which is also used in some power electronics what are those called? Contactors, okay. So there (())(17:32) system if you go and look at in the power system we are using now S F 6 discharge, okay. Okay this is just additional information. Now no one is using Freon but I think many VLSI companies they are still using Freon and not telling probably, okay.

These equations are clear. Is that two mechanisms are clear? Is that two mechanism clear in R I E? One bombardment, other chemical reaction. So it is not non reactive, it is a reactive specie of fluorine can be used. No is not there. But it need not be fluorine, any other species. Now this is a isotropic etching. Why it is called isotropic? Because it is not directional or it can etch any direction, it does not matter. I want to remove every place anyway and it is used in ashing of resist using oxygen plasma.

So where is the reactor which is called the very famous reactor? Actually that is how we started with. It is called barrel reactor. This is the end view and this is the side view of the reactor.

(Refer Slide Time: 18:42)



You can see from here there is a boat on which wafers are kept. Then between these two plates we apply R F. This is one plate, this is your other plate anode cathode. This is the R F applied. I enter the gas. The most important difference before you draw figure you see this. There is a small some wired mesh I kept there, is that clear? This is a wired mesh, okay, which is essentially is grounded not really shown correctly. This shield is it is called shield. So the shield is all around the wafers. That means shield is all around the wafers.

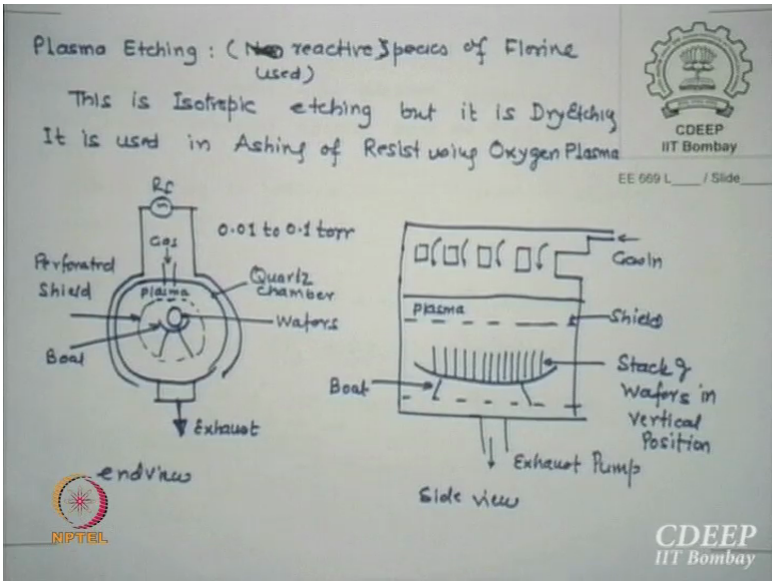
This is the shield so wafer sit inside the shield and shield is grounded. So when you get a discharge, ions will be created which are positive plus charged ions. How about this when they pass through shield they will lose their charge to the shield. So what is left there is only say oxygen plasma is used so only oxygen atoms which are not necessarily charged. Some will still pass through but few of them. Most of them will be only oxygen and this is O. Why O? Because the reaction I am using in the plasma I am not getting O 2.

So what is the difference between O 2 and O? Nascent oxygen is highly reactive because that does not get a bond so it starts reacting. That oxygen then reacts with any photo resist carbon material and will convert it to carbon dioxide, okay, which is exhausted out and the whole wafer is cleaned out, okay. So these etchers are only used for photo resist ashing or something where you want to remove all the films across, okay. So this is another etching which is used.

This is dry etching, the other is also dry etching but what is the difference there? It is reactive as well. In this case it is only reactive but it is not bombardment. Is that correct? Reactive ion.

Here only reactive species. Is that correct? That is the only difference between the two etch systems, okay.

(Refer Slide Time: 20:56)

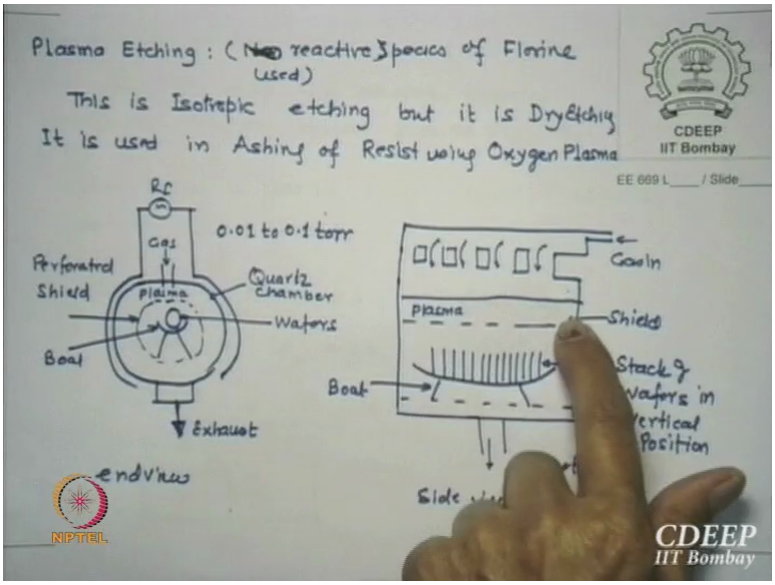


The first one we said it has two possibilities. One is ion bombardment, other is etching. So I want to create a model based on this mechanism. For R I E there are two possible mechanisms etching. One is I said related to chemical etching, the other is ion bombardment. Is that okay? Figure is drawn. This figure is also so very popular. Also you can go to applied materials lab. They have one barrel etcher kept in corner.

So there are two possible models are created for describing the etching in plasma systems, okay. Which one? Oh, this is R F coiled. Also know there is a called diffuser. Sorry this may cause a little ahead but below there is a diffuser. That means there is a circular plate and holes in. This is called diffuser. Why? Because gas if it comes from one end it will be heavily pushing on one end only. So it actually allows to gas go through number of holes so that everywhere plasma pressure is constant.

Is that clear? If I push gas from all sides, everywhere the plasma pressure is uniform. So that is why it is but just above this is the R F coil just above that because you know otherwise the gas which is coming here, this plasma will be less ions than this plasma, okay. So I want uniform plasma.

(Refer Slide Time: 22:32)



So I want to push gas everywhere. Is that clear? The two models which allows you to model the etching is linear etch model and saturation absorption model for R I E. It is called R I E. The first model, both are for R I E. I think I have made a mistake. But linear etch model says it (assumps) assume something. Yes we have say two mechanism, chemical and ionic. However our assumption is during the etching they independently react. They do not interact with each other.

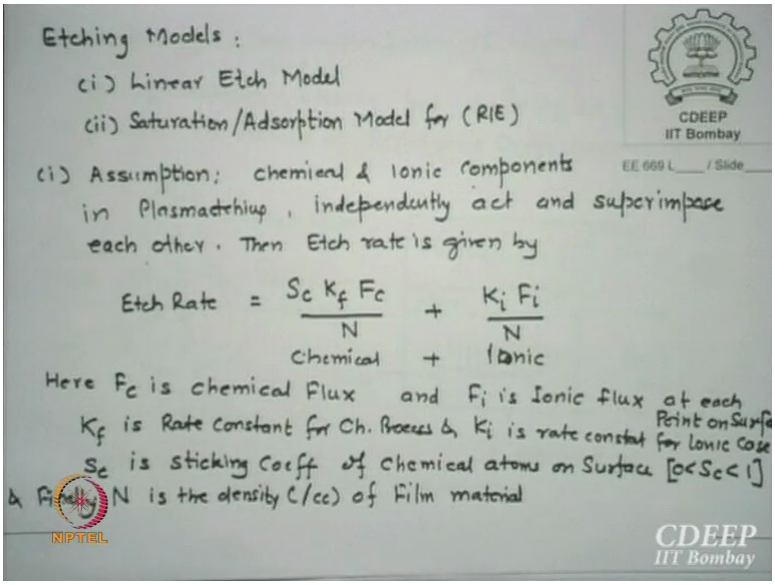
One is happening, other is also happening and they are like a superposition, okay. So we say chemical and ionic components in plasma etching independently act and superimpose in their result. Then the etch rate can be given as one due to the chemical reaction and one due to the ionic reaction or ionic bombardment. So we figured out this model is given and can be derived also but I just wrote down. There is a term called S c K f F c by N plus K i F i by N where F stand for flux in all cases, okay.

So F c is the available gas flux. What is flux is? Number of atoms per unit area per unit time. Which is decided by what? The flux is decided by the pressure at the gas inside, okay. So I know F c how much gas pressure I will use. I know fluxes of these. Here it is essentially is called rate reaction constant due to chemistry whatever is the reaction rate between the specie as well as the substrate which you are etching. S c is called sticking coefficient, okay. We will discuss this little later.

And N is the number of atoms or number density per cc of the film. So whichever film are there but like in silicon what is the N number? 5 into 10 to the power 22 per cc. For different

materials this number will be different. And looking at this side for ionic, K i is the reaction rate due to ion flux, rate constant. And F i is the ionic flux at each point on the surface.

(Refer Slide Time: 24:58)



Now why this word S c has been introduced in the first case? Because the chemical reaction can only take place if the incoming atom actually stays on the surface. You know it has to react so it must stay there, okay, because it will stay and will start moving. And then can may go emit also, okay. So as long as it sticks there only then it can etch. So the coefficient which describes this sticking possibility which overcomes the surface energy is called sticking coefficient, okay. For different materials and different gases, S c are different.

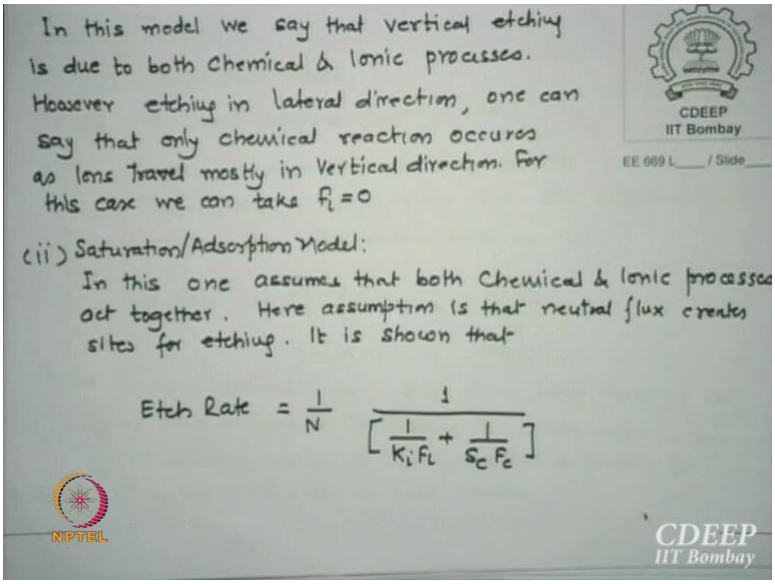
Typically S c all item getting attached when S c is 1, none is getting attached when S c is zero. So S c remember is typically between zero and 1, can be as high as point 5 in most cases, never 1. Not all atom stick, partly they move and emit also, okay. So S c is not more than point 5 but certainly it is not very small number either, okay. So if I know now F c, this and I know K i for the film I am etching or the flux I am using then I know my etch rate. I know my etch rate due to what? R I E reactive as well as bombardment total etch rate is this.

If I multiply it by time then what I will get is film thickness, okay. So if case of film thickness I want to etch in certain time I know how much should be etch rate and depending on the etch rate I can find what fluxes I should use or pressures I should keep so that this film can be etched in this much time. Is that clear? This is how back calculations are performed. First you are asked A f should be close to 1. This I want, this I want. So you have to back calculate and come and find how much should be the gas pressure, okay.

That is how we adjust the gas pressure. What is the other method? Keep on changing the wall, it may take 100 readings. Someday it will hit correct, okay. Okay, so please remember vertical etching is in which direction? Vertical etching is because of which mechanism? Both chemical as well as ionic. But lateral will be because of what? Only chemical, of course few ions do scatter but generally most of it will do chemical etching.

In this model we say vertical etching is due to only chemical and ionic processes however etching in lateral direction one can say that only chemical reaction occurs and ions travel mostly in vertical direction. And in that case to get a lateral etching what should I do etch rate? Put F i equal to zero, no flux due to ions in the lateral side. So I can get lateral. Is that clear? How do get lateral etch rate? Just make F i equal to zero, okay.

(Refer Slide Time: 28:01)



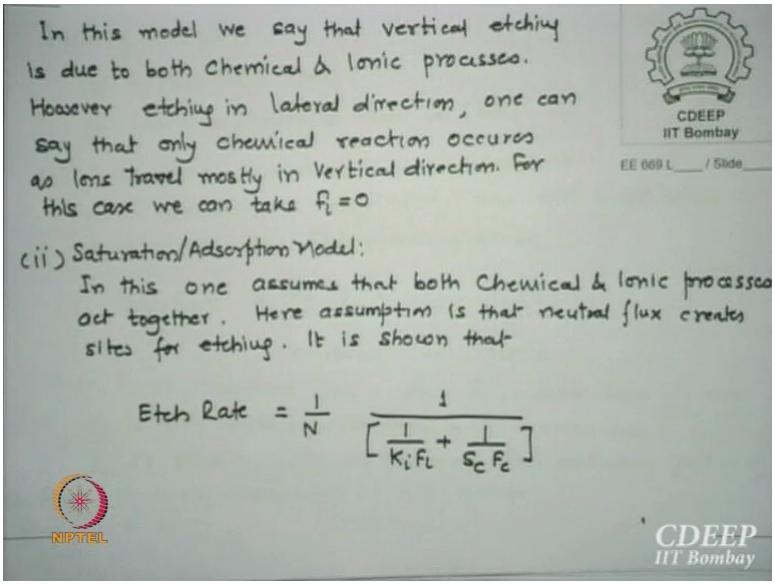
This was one model suggested which worked reasonably well in many cases but other assumption was both processes are independent itself is not correct, okay. Firstly because when the ions come, our assumption is when the path is sufficient they do not scatter. But it does not happen. They hit the walls and they actually release electron which actually takes the energy from the ions so the direction changes. So at the end one cannot say guarantee that everything is coming only down.

Some ionic concentration may go on the lateral side as well because of both acting together and affecting each other, okay. This is our first assumption we said but in reality it may not happen. If that happens that both are acting together, one can derive little similar fashion. Etch rate is 1 upon N, 1 upon K i F i plus 1 upon S c F c. And this model has fitted well in

some gases particularly flowing based but did not fit in the other gases etchings. So one does not know whether this model is correct, this model is correct.

Whichever fitted to your data you may say this model in our case is correct. So it is a fitting data model which we did enough etchings and we figure out typically which model normally reacts, okay. But this has more chances of fitting. Why? Because it is assuming both together, okay. Reaction on each other on both sides. So there is a possibility but modelling a scattered ions is not very easy, scattering ions, okay.

(Refer Slide Time: 29:43)



So this becomes only first order equivalence and therefore not very accurate but otherwise this at least takes care of interactions, okay. So this finishes the possibilities of etching any film. What was the etching we did? First we did wet etching and then we said dry etching. So why do you want to prefer dry etching? Selective etching and anisotropic etching. Why selective? Because one finds that wet etchings can lift many things. What happens?

If there is a field and there is a window, the solution goes in and from the corner it may go in, okay. And then lift the upper film which is stopping actually. So the worries are in wet etching is solutions are not that socialist in their thinking. They do not think that this I have to go I have not to go. They go everywhere. Whereas these are more capitalistic people, they go vertically or wherever they want and therefore they are more accurate, okay. And much more less clumsier films than anything.

However what is the problem with ion etchings? Ions are sufficient energy so they are hitting what? During etching the lower surface will be damaged surface so you will have to recover

that damage. So you have to (())(31:13). That means one thermal budget you are increased. Is that clear? So it is not at free cost. Mostly you have to take it away, clean it, redo it this and again you have to put extra thermal budget for actually getting this better film etching. Is that okay? So nothing at zero cost, okay.

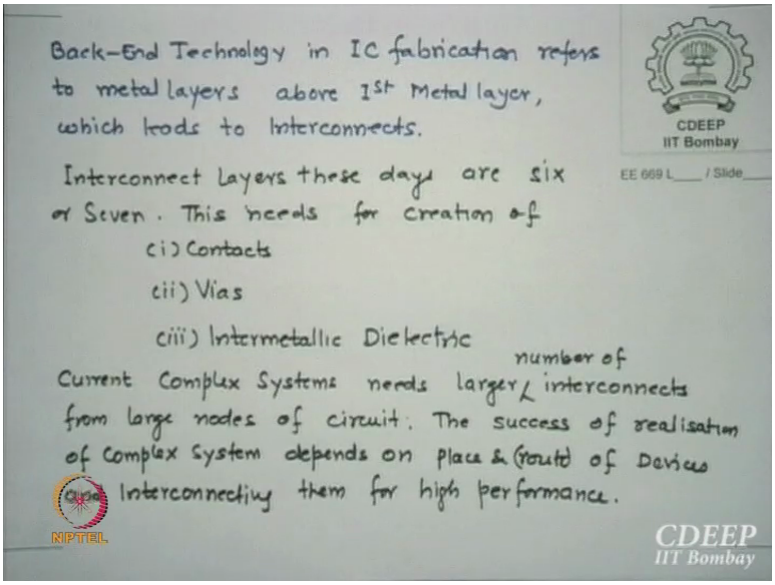
So this finishes etching. The last part of this course which should have taken more efforts than what I am doing today, as I said the back end engineers are the ones who make the chip success, system success and therefore should be really looked into much deeper. But since they overlap with the designers so one thinks whether designer should do this part or the technology part? So I am only going to talk about that part in which technology effort is required.

The designers are the ones who decide what I should do, okay. So what is back end? Essentially back end means the first one or two layers of metal which are from the devices or to first interconnect as we call, may be part of the front end process, okay. Anything above this is as number of layers you create these are called back end technologies. Why? Because they decide the interconnect for the chip. Typically how much is the length of the interconnect on a good Intel chip?

I said you 5 kilometres on 1 point 7 centimetres sized chip, okay. So one can think the one which is going to decide everything for you I am finishing in 10 - 20 minutes, okay. Though actually it should demand that at least 8 - 6 hours I should spend only on back end because these are the ones which will make chip failed. Everything goes well till front end and chip fails at the back end, okay. So issues which are very worrying are back end issues but as I say I am not going to spend time though I wish I could.

So back end technology in IC fabrication refers to metal layers above first metal layer which leads to interconnects. And I said these days six or seven layers of interconnects are, even 9 have been tried now for a more complex circuit. So these are essentially what? So what are the problems in creating number of interconnecting metal layers? I have to talk about contacts. I must talk about vias through which something goes down and of course then between the two metal layers I need a dielectric to separate the metal layers interconnect lines.

(Refer Slide Time: 33:54)



So when I say I am looking for back end I am looking for contacts, I am looking for vias and I am looking for intermetallic dialectics. Of course route is wrong. So how do you place devices or transistors and how do we interconnect is another game. Typically if you are the circuit men they will say they should be something called local, something called global interconnect. The local interconnects are close to the device source rates. These are called local interconnects. Anything away outside that is called global interconnect.

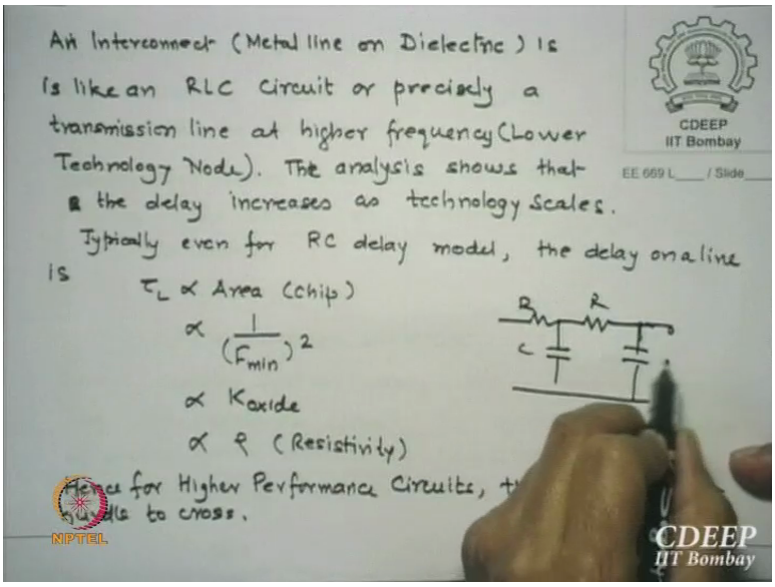
Now we keep saying some should be synchronous to the clocks, some should be non synchronous to the clock, okay. So these are issues which circuit people are worried about because to maintain certain speed of a particular processor or any circuit these are called globally asynchronous and local synchronous is what is one tries in most cases, gals is the technology which is used globally asynchronous. Please remember in circuits, asynchronous does not mean clock zero. It is not synchronous with the system clock, okay.

So interconnects run at different frequencies and actual device runs first interconnect runs with some other frequency which is synchronous to the actual system clock. So please remember where do you keep your devices and how do you interconnect on so many layers? This another software possible is called place and route. But with many layers of metal it become very difficult for even place and route to get smallest amount of delay which is what we are really looking for. There can be multiple choices, okay.

So here is how we start interacting with designers. Now an interconnect which is a metal line on a dielectric is like an R L C circuit. Typically we always believe it is R C circuit but with

great amount of research we did for many years that was sufficient. That is typically around few gigahertz a metal line on a dielectric can be thought as R L C. However what is I am saying? This is only R C transmission line.

(Refer Slide Time: 36:24)



So most of the time we calculated using simple (())(36:30) formula or others or actually solving the transmission line, the delay of this line and we say that is the delay we are worried about, okay. However now the metal itself is becoming thinner because of scaling and the frequencies are increasing. So, simple wire is not a resistor alone, okay. It may actually have now inductor along with the resistor. So circuit of a transmission line is now more like R L C.

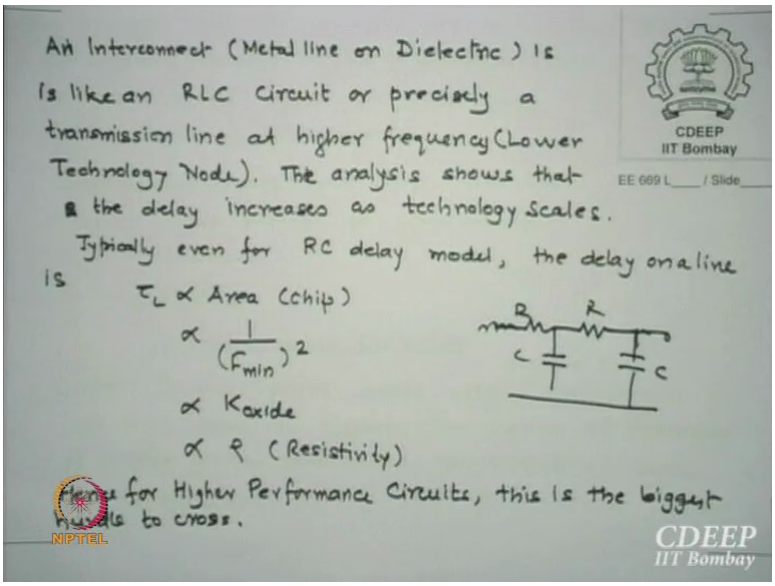
There is also R across C now because of the capacitors are not ideal capacitors. So the kind of lossy transmission line networks solving is what is now needed if I want to find delay in interconnect. Now this delay has a problem. If you look at a very simple R C model, larger the length of interconnect I run larger is the R. C maybe area wise, R also is also increasing. So if I take a signal from one and two other in a longer line, the R C delay will be larger.

And if this 1 upon R C is the same as your system clock or close to that, 1 may actually become zero by that actually clock age, okay, equivalently saying because you are getting one phase out because of your delay of one cycle. This essentially means your whole data transfer will be actually (())(38:09). So we must see to it that this R C delay or R L C delay should be much smaller than the actual system clock I am running, okay. That is why that asynchronous, we do not want that all the higher frequencies.

That is why globally we want to asynchronous. So this idea that R C should be reduced, essentially if you look at the time constant associated with delay tau L, larger the area larger is the tau L. Smallest the feature size, if the feature size is small it is features size square inversely it goes to delay.

So if you are working on 11 hour metres actually the R C delay will be much longer, okay. Because everything is thinning down, okay. So actually delay becomes worsen if I go from 45 to 30, 30 to 22 to 16, 11, the delay is even worse. So I will have to do additional features to see that delay is avoided.

(Refer Slide Time: 39:16)



It also depends upon intermetallic distance and the dielectric you keep so K ox. And it also depends on the metal resistivity. So should I use aluminium at all now? So I may use for which other smaller resistivity so copper form example. So these are the area, when I look for higher performance circuit I must see to it that I confirm to a circuit performance. Why I am worried about? Because device people always believe that they have made source drain, they have made channel lengths.

So what? Everything what you wanted I have given you, okay. That is what the front end people feel, okay. But the problem starts afterwards. The circuit does not work. Each start blaming this was your problem, this was my problem. But then the person who put the money, he lost it. So he is more worried. So, one issue which you must not forget that the back end design is very crucial. So how do I put interconnect? I suggest technology which metals I should need.

What is the delay maximum I expect from you, okay? All these issues of course if you do not do it I will put buffers after every small length of interconnect. I will put a buffer which will reduce my delay, okay. However buffer means I consume power, I put extra area and everywhere I have no buffers. So of course circuit person will say I will get you out, okay. But that is not the best solution. Large size and it is called non inverting inverter. In the sense epsilon by this C is proportional to K, okay.

But we are not going for gale. That is what I am trying to say. All this time I said you Ramayana, now do not go Ram and Sita, okay. They are outside device now, okay. We are only in the interconnects, okay. The dielectric between two metal lines need not be high K. If that happens it is the worst thing to happen. Oxide does not mean silicon dioxide, insulator in general. For all these years here now the first layer still can be aluminium as a contact and as an interconnect.

What is the difference between contact and interconnect? Contact is with the source drain, okay. So it can be a contact material but the line running need not be aluminium. So but earlier we used to run aluminium as an interconnect as well as for the contact. Why aluminium was so preferred material? There are many reasons why we started with aluminium. First of course if you only look for conductivity then aluminium has delay 10 to the power minus 5 (())(42:18) per centimetre is the conductivity.

So very low. So what is wrong with aluminium? Fantastic, aluminium has all good features, okay, which one wanted but it is still not the best conductor. Copper is still better conductivity than aluminium. Gold and silver are also equally good but somehow in India or abroad anywhere if you talk about gold everyone fears as if it is very costly. In fact platinum which we use very often is much costlier than gold. But when it comes to gold is very costly. So the gold technology is never preferred for not being costly.

Gold is very bad centre creator in silicon. It gives (())(43:01) means one level is above the lower which should be opposite side. The acceptor is above the donor which gives huge amount of traffic, okay. And that is essentially killing the lifetime. That is what use actually for killing the lifetime. So gold I do not want to come. Silver get oxidized without doing anything it just see this and blackens, okay. So left is aluminium. So people worked on aluminium. Then why not copper? Copper also get oxidized, one is that.

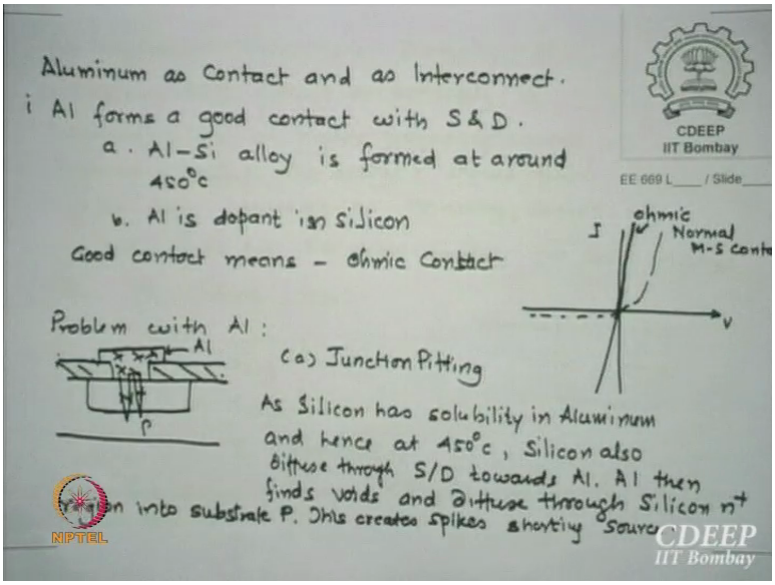
Secondly copper also has a trap level in the silicon, okay. So my time we never used a copper line of gases which is very good to run. Then all gases should pass through good conducting gas lines but we never use it because copper was anathema. It may kill all my devices, okay. May or may not happen but we always were worried that copper should not be used within 100 metres of the lab. No copper tubing, okay. Now but that is the fantastic part. Aluminium only top layer it forms but the lower layer it does not form. That is called anodization.

That it protects the aluminium is only thing which did that whereas copper oxide it actually porous down to all the films. You have a point. Aluminium has all the features which you wanted actually. Is that clear? So aluminium has all good features. Why we went from aluminium was that R C delay. Someone said R . Then we started looking for better conductivity material, okay. That is the reason.

Typically aluminium has many advantage of silicon technology. Aluminium silicon alloy is formed around 450 degrees. It is called (())(44:54). So aluminium silicon forms alloy is around 450 degree. Even lower 423 itself, but 450 most of it forms alloy. Also aluminium is a dopant in silicon type 3, okay. So at least on P it will make P plus so good omissity it can create, okay. It is fantastic. It has a good very low temperature it can form a contact alloy between silicon and aluminium. Please remember alloy means better contact same uniformity.

So this is something contact I should say if a normal metal semiconductor junction which is called short barrier. If I plot I V, this dotted line is normal metal semiconductor (barr) short barriers. Whereas the kind of contact I am looking or ohmic contact I am looking should be very low resistance both in forward bias as well as in reverse bias. So almost ideally I want like this, zero resistance. But that is not possible so close to that, as low as possible.

(Refer Slide Time: 46:04)



This is what I am trying to achieve. So silicon and aluminium forms a good alloy. Aluminium is a type 3 dopant, so it will give even better contact resistance or lower contact resistance and therefore more likely to give me more ohmic contact. So no voltage drop at these contacts. What is the problem if the voltage drop goes on the contact? If source drain has additional voltage drops so the delay increases because for the channel V D A is decreases, okay. So I am trying to see that my source drain contacts have least voltage drops.

Source drain itself will have some drop but this is additional. When that there is an area with it the line may run longer, much more area may create. So I somehow wish to avoid this part very clearly. So I want to make good omissity, okay. However aluminium has its own problems. One of the major worries with aluminium as we have found was when I made a aluminium contact to source on drain, for this is my aluminium and this is my N silicon or whatever it is.

Please remember silicon has also solubility in aluminium and aluminium has a solubility in silicon. Is that clear? This is called face diagram. Aluminium dissolves in silicon and silicon dissolves in aluminium. So around 450 itself silicon has sufficient diffusivity in aluminium. So what happens when you start (())(47:46) for allowing the contact, part of the silicon actually comes into aluminium, okay. It creates voids because silicon gone from there.

Aluminium tries to defuse down because it find voids and silicon down so it comes down. So now aluminium contact to this substrate. This is called junction pitting. So why? Source drain gone, is that clear? The top contact is to the substrate, no source drain. Of course this is not

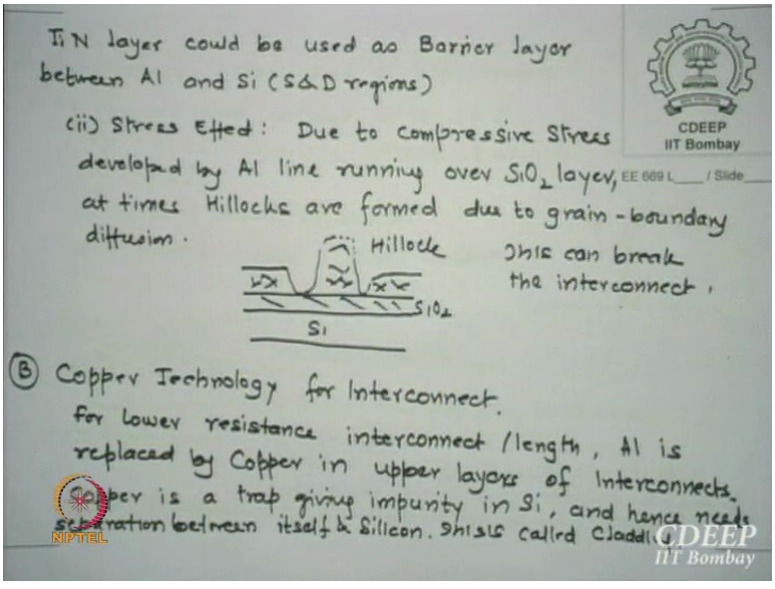
that everywhere it will happen but additional resistance certainly will occur and there will be an issue and sometime circuit may not even function. Is that clear? So the worries are of junction pitting is huge. Actual when it will not have occurred if source drains are deeper.

In my time we had 3 microns source drains. Aluminium film was hardly 1000 to 2000 angstroms. So that was very unlikely that it will go to 3 micron deep. So I may never saw junction pitting in earlier times but when our source drain become 1000 angstroms then this pitting is very possible, okay. So you can see scaling has actually created (())(49:00) issue. It is for the pitting problems. 450 has to be done to make an alloy between aluminium and silicon.

This is your tactic. If I come from aluminium this to silicon this, at 450 they will form alloy. Look for a face diagram, okay. So, is that point clear? So first major worry in a back end is that if you make aluminium (())(49:29) scale down devices this issue may actually hurt you, okay. So what should I do? Between aluminium and silicon I should put something in between which will not allow aluminium or silicon to cross diffuse. That is called barrier. Titanium nitride is a good barrier layer, okay.

But it is certainly not as good a contact to silicon nor as good a contact to aluminium. So normally thin titanium nitrate layer is essentially interposed between silicon and aluminium layer just to make it barrier for inter diffusions, okay.

(Refer Slide Time: 50:21)



It is also called buffer layer in some books. But most VLSI people call it barrier. If you have noted down this is another effect. Of course there is third effect probably I will say. The

present date tendencies to actually give some stresses to the silicon device, what for we want to stress the device? Anyone? Mobility enhancement by compressive and by tensile one of the mobility could be enhanced. So we are now really trying to shrink the device and trying to even create resist, okay.

Now this compressive stress particular is very harmful for aluminium. If aluminium line is running and this structure is similar shown here and it receives because the device is under compressive stress, part of the aluminium actually. Please remember aluminium is a metal, it is not crystalline. What is the form of metal? All metals are amorphous or polycrystalline, at best poly or mostly amorphous, okay. Now there is already drain boundaries in the polycrystallines.

Now what happens when the metal is running over to layer due to the drain boundary diffusions or movements some aluminium moves away from its position and actually climbs one over and actually break the contact between . So it is a Hillock formation which actually one can see under an SCM. It is not a very, even normal 100 x to 10 x microscope you can see this Hillock, okay, and a breakage of contacts. The similar effect if not same is also seen by electromigrations.

What is electromigration I said? If in the metal layer huge current density is applied, 10 to power 5 amp per centimetre square then the number of electrons ions are very large there per unit area. This is called electron wind. Electrons are more mobile so they try to move away and this movement of electrons causes a force. And if aluminium ions or atoms there cannot stand to this force then they jump, okay. Similar like they jump and that is called electromigration.

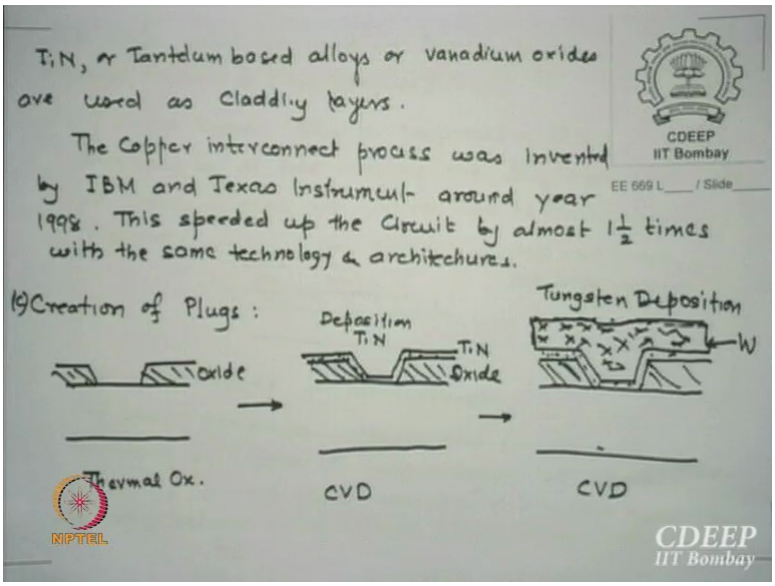
If you put titanium nitride that itself will stop it because the electromigration coefficient of nitrite is much lower compared to what is for aluminium, okay. So copper is the best for electromigration. It is not easy, copper is more metallic than aluminium in drain boundary system. So it does not climb so easily. So higher currents can easily be taken by copper compared to aluminium. Copper has better conductivity.

So we start looking for this. To give some history on this, the copper technology was first tried by Texas Instrument and of course there is a word for it. Maybe I will come back to it. For a lower resistance interconnect per unit length aluminium is replaced now by copper in upper layers at least. Copper is a trap giving impurity in silicon I just said, okay. And hence

needs again separation between anything to silicon it must get separated. So what should I need between copper and this?

Some kind of a barrier and that wherever copper is going it should be surrounded by that barrier material. So what is surrounding system should be called? In fibre what do we call? Cladding. So all copper lines or copper this must be cladded by titanium nitride or some other barrier material, okay. So titanium based alloys or aluminium based oxides, nitrite sometimes both or titanium nitrides are excellent cladding materials and they are normally surrounding copper lines, okay.

(Refer Slide Time: 54:48)



Just to give history, copper interconnect process was claimed by, I do not know whether it was invented by they claimed they have patents, both by IBM and Texas Instrument around the year 98 and the interesting part why they did it? That by just replacing aluminium by copper at least the three upper layer interconnects, they found their circuit started speeding up by one and a half times without architectural change, no dimensions on anything. Just putting copper layers they figured out the speed to one and half times.

So 3 gigahertz processor or those of course 2 point 1 gigahertz suddenly became 3 gigahertz processor for doing nothing except this new process, okay. So that is why copper has replaced at least for top interconnects. That is decided by the technology films. Films are decided by technology people. Typically vias are of the order of 1000 angstroms. So only that much copper lines can run. No one it is (())(56:00). No , it is not circular.

I showed you a plug. Let us say I want between the two metal layers connection, okay. So I did a hole through between inter oxide and connect top portion to the lower. But can I connect the same upper metal down? So that I do not want to do. I want guaranteed contact from upper layer to the lower layer. So I create a vias. Which process will create vias? R I E, because I want vertically dimensioned etch, okay.

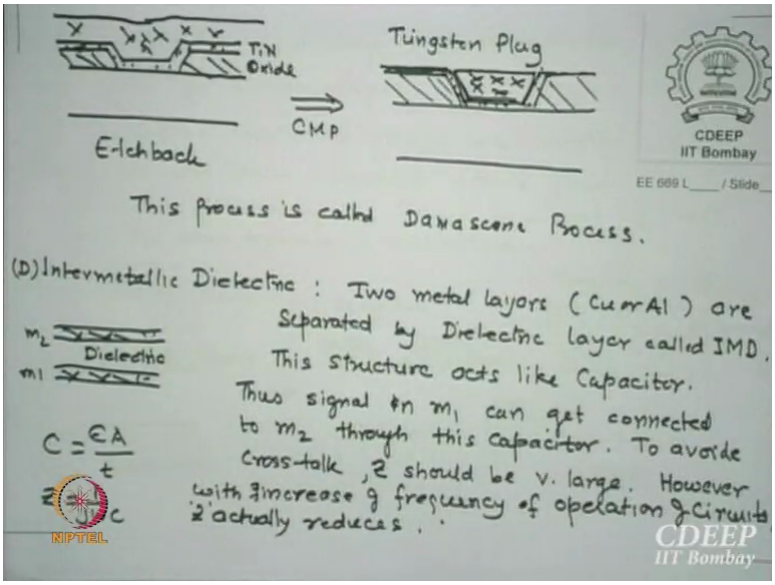
Then I will put some material in between which is conducting and easy to flow in this so that it fills up that vias, okay. This is called plug which fills up that vias. So how do I create a plug? Let us say I have an oxide. I open a window. Then I deposit titanium nitride first. I open a window by lithography and I actually create titanium nitride, okay. Film on that which is my C V D and then I deposit tungsten either by C V D or even by sputtering but normally by C V D.

Yes, the layer is the gas which is used tungsten hexafluoride and that gas is passed because in C V D what is the thing I kept on saying? It is it will follow all the contours of the wafer conformal therefore if this is like this, the tungsten will also flow this, okay. So I deposited tungsten sufficiently thicker than the vias itself. So you can see from here, the thicker metal was deposited, okay, in the all portions other than this, okay. Then this thickness you know many time the upper layer is deposited by sputtering.

The first you do C V D and then just deposit dump by sputtering. So that I do not have to worry about doing. The one sheet has come up, I deposit even thicker on that, okay. Is that okay? 1, 2, 3. Then I do little etching of the tungsten. It is called etch back. This is very important process. How much to reduce tungsten? Okay, because the next process which we are going to do are chemical mechanical polish and very thick layers cannot be polished out. So I have to reduce that tungsten thickness, okay.

Earlier I did not know I put sputtering, I put some large number. So then I monitored it and say okay at least 80 percent I must get rid of that. So that is I normally do by etching only, okay. Since tungsten is everywhere it etches everywhere. It reduces the thickness of tungsten and then I do chemical mechanical polish on this surface and tungsten just gets inside.

(Refer Slide Time: 59:18)



And please remember your nitride will be removed by mechanical this so there is a titanium nitride, in between there is tungsten block or plug as it is called, okay. This process is actually called Damascene process. Is that okay? How do I do it? Etch the silicon dioxide, titanium nitride, tungsten thicker, reduce thickness, do C and P to flatten it, okay. So now I have a tungsten plug which is created here. The next problem as I said, vias of course as I said by ion etching, contacts I did.

I now want to see how between the two metal layers which dielectric I should use, okay. Two metal layers whether it is copper or aluminium are separated through a dielectric layer and called I M D. Why I M D? Intermetallic dielectric, I M D. If you look at an IMD structure here this is the metal, metal and a dielectric. What it looks like? A capacitor, okay. So now the problem starts for me. The structure is now a capacitor.

So if any signal is passing on any of this metal line, let us say m 1 is carrying a signal and let us say the capacitance of this is such that the z is 1 upon j omega C, okay. If C is large enough then z will be smaller. If omega is large enough, z will be smaller. So high frequency signal and with larger C if we have, there is a connection at least at that frequency connection between metal 1 and metal 2. This is called cross talk, okay. This is called cross talk.

So what is the criteria I should to avoid cross talk? There are many things we can try. Some are see to it that the z between the two metal is very high at that frequency of operation. Now to make z very high another problem which I am worried about, if I want to make z high C should be smaller. If C should be smaller that means t should be higher. But if I use t very

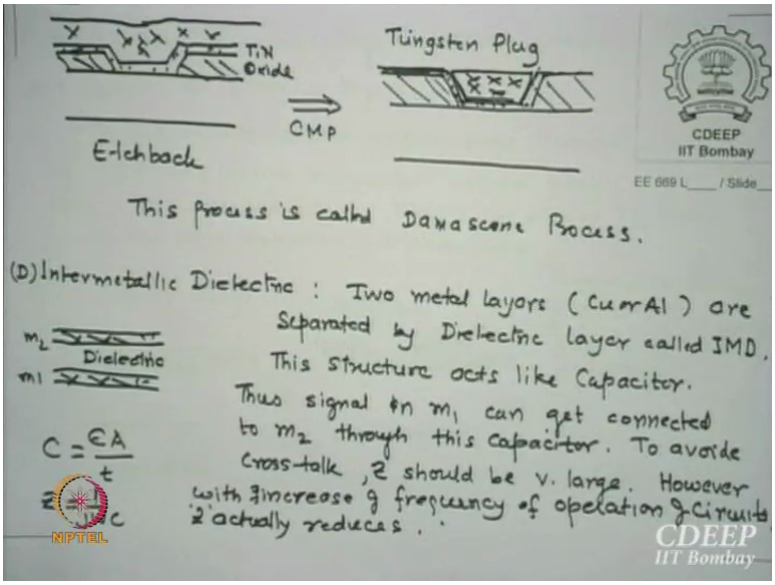
high what will happen? The vias connection from top to bottom I cannot guarantee vertically down all the way deeper vias, unlikely to get same size vias, okay.

It will actually get weaker on that, okay. The way it happens. If thickness is larger I need deeper vias so it is reliability issue. Tungsten though they said that can go any conformality but if the vias size becomes smaller in the lower side, gas just goes above. This causes surface tension effect. So it does not cross the surface energy there. So it does not stick to the lower side. Is that clear? So there is air gap. Is that point clear what I said?

If the vias size is smaller, the gas goes in but does not touch the lower surface because there is what is called a surface tension effect which means there is no surface waiting. So the gas just goes above and does not connect with the lower end which means there is a air gap that means resistance of vias is going to increase now because how much is air gap it depends on that, okay. Now this issue was very tough when I started shrinking.

So obviously I do not want to increase thickness. I wish I could but I do not want to. A is anyway increasing drastically. Why? Because interconnects are lengthening and lengthening so area of the capacitor is every time increasing. So smaller the dimension device I need, larger interconnect lengths I am using because I want larger circuit. So more capacitance I am creating out of A. T I cannot increase, A is going to increase anyway. So see the worried I am. So the only possibility left from here is to reduce epsilon. Is that correct?

(Refer Slide Time: 01:04:03)



So to avoid some kind of this, increase some thickness and increase intermetallic dielectric whose epsilon is smaller. Is that point clear why back end people are worried by this?

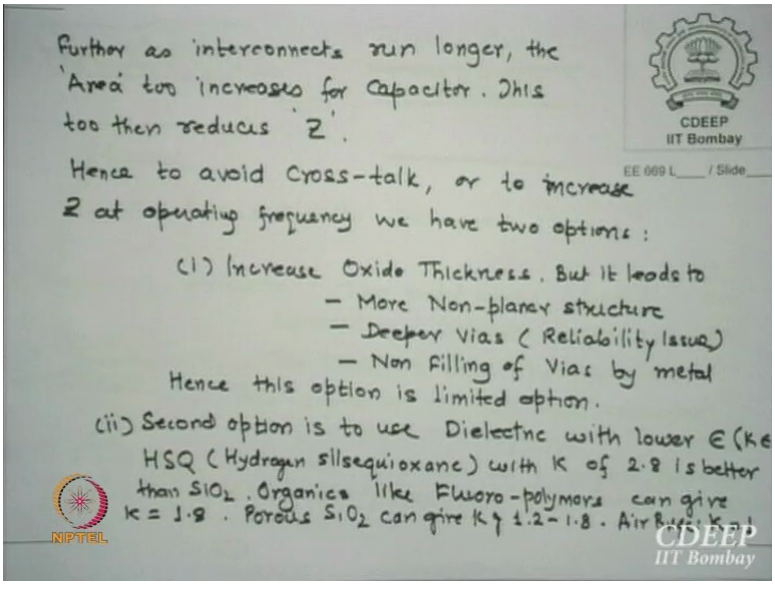
Because they find cross talk too high, okay. Is that okay? Function wise I want to make z infinite or z larger, C should be smaller, A is increasing, t cannot be reduced too much so epsilon is the only which also I cannot play too much. What is the smallest epsilon I can get? 1. What is that? Air. So air bridge is now being tried.

You are right. Further as interconnects run longer area too increase for capacitor, this too then reduces z. Hence to avoid cross talk or to increase z at operating frequency we have only two options. Increase oxide thickness but it lead to a more non planar structures. Deeper vias reliability issues. Non filling of vias by metal hence this option is of limited option. It is an option but limited. Second option is to use dielectric with lower epsilon, K epsilon zero.

So one of the material tried was hydrogen silsesquioxane which is H S Q in short which have the dielectric constant of 2 point 8. How much is dielectric constant of silicon dioxide? 3 point 9. Nitride, 7 even higher.

So I want to reduce. 2 point 8 is at least lower than silicon dioxide so I may use this kind of a glass which is a organic glass which I can put which will slightly improve the capacitance ratios and therefore some better cross talk property. Not good, better and not best I want but something better.

(Refer Slide Time: 01:06:21)



There are also now people are trying organics like fluoropolymers as it is called. These are polymers with fluorine. Their formulas are too big a carbon chain. So I thought just to show on a page lot of carbon rings does not make sense. So I did not draw but it has 12 carbon chain, two in the one spin side, four in the other spin side. So lot of chemistry involved. All

sort and done its dielectric constant is 1 point 8. So even better than silicon dioxide, better than H S Q.

So there is another material similar like polymer. It is called polyimide which has the dielectric constant of 2 point 1. So first we tried H S Q, then we go to polyimide and now we are looking for fluoropolymers for the inter metal. So what is the problem with these dielectrics? For a higher temperature and higher humidity their properties change so their protection is very difficult so they need huge (())(01:07:29) on the top, okay. H S Q flows but its flowing temperature is less than glass.

It is 600 or something it flows. But the way it is it is sol gel deposition. So it is a gel form. So you actually dip and spin. That is how it is. All processes normally for polymers are solution gels so you put drops and spin, okay. That is how it is deposited. There is another material people are recently trying and to some extent successful. In 1992 one of my Ph d student worked on porous silicon now S i O 2, porous silicon.

This is also porous as well as porous silicon both are tried. What is porous silicon anyone? It is not an amorphous gas silicon or amorphous but there are pores along the crystalline. There is no actual crystal anywhere. There is no order and there is gap. These are called pores. It has a very different property than, some other day. That time we did not realise that this can be useful in VLSI. We were looking for LEDs because it will give electroluminescence. Porous silicon is used for electroluminescence, okay.

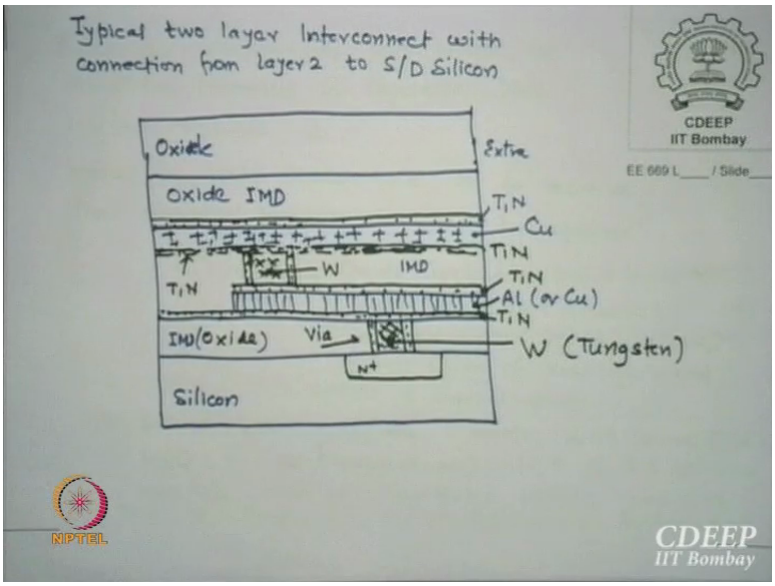
So we are looking for LED replacement from the semiconductor. So other normal so we were trying for porous silicon. But we only left doing a lot of theory and never made a device. So I do not know whether it gives light. Hopefully yes. So no, the silicon is etched in a particular solution so it gives lot of porosity deep into the solutions. Some other time maybe I will discuss with you.

Typical porous silicon and silicon dioxide layers have a dielectric constant around 1 point 2 to 1 point 8 which is what the best probably people will be able to achieve now, okay. Of course last one is air bridge. There is a method of creating pressured structure that is shield structure in which air can be filled in between the two layers. This is called pressurized air bridges. It is a very complicated technology. It is done under very high vacuum 10 to the power minus 16 but it is very difficult to maintain.

If small leak occurs it will just go away but it is possible to trap air in a smaller region, okay. That is done through MBA machines. That is done and it has a K of 1 ideal so between two metal layer only air gap. It is doable but it is very complicated, costly as if not tried. No, between two metal layers when I actually I put a air bridge there and then I put the next metal that is why I say MBA machines so that there is a very tough situation. But it is doable. Some effort has been successful but not hundred percent so I just do not want to discuss, no one in industry as well.

Last but not the least with this course is this. This is another back end problem. From one metal layer to the other metal layer how do I go interconnects? That is most important. You have 9 layers, 7 layers at least but 9 also can happen. So from the lower layer to upper layer or upper layer to lower layer how do I come down? So here is the method of interconnection between let us say this is my source drain, one of them, this is my silicon, this is my N plus, then this is my IMD which is there. On IMD I etch this window. Please look at it.

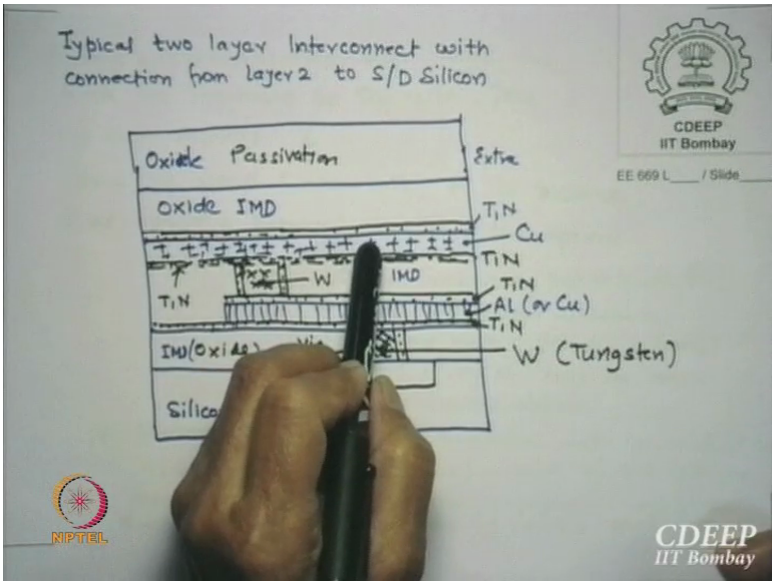
(Refer Slide Time: 01:11:29)



First draw it then we will discuss, okay. We have the silicon, we have a source drain N plus region, I have deposit IMD on that, open a window this much, deposit titanium nitride, okay. All two sides, then deposit metal which is tungsten in our case either by first CVD and then by dumping and then what we are do? I do a CNP so that I get a plug. Then I deposit titanium nitride once again. I deposit the next layer of metal which can be aluminium or copper now. Then over which I have another titanium nitride.

Please remember any metal layer should be barrier by both sides by a titanium nitride. This so what you are saying vanadium plugs this occurs here exactly same as titanium nitride. Then next plug which you want to create should not be in this direction because this oxide if you actually open here there is a misalignment issues because you will never see this if you actually put the next contact on that.

(Refer Slide Time: 01:12:56)



Now I can see this in my pattern and I can start looking this size where I am putting the next contact. If I use it here one is porosity issue, some materials can go through. Second this contact I will not see below, okay. Because lower side will be mask there. So if I separate them first thing I must separate them. So our next plug is again titanium nitride, tungsten plug CNP. Another titanium nitride, another metal layer, another titanium nitride then the IMD in case you want to go ahead or if not you put the final passivation oxide.

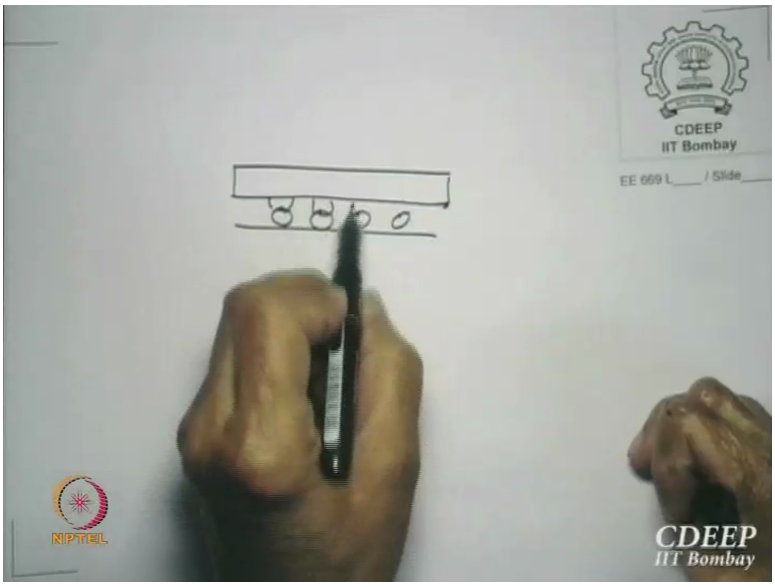
Actually it is mostly not oxide either it is borosilicate or glass or S i 3 N 4. The top layer of a silicon wafer is protected by silicon nitride. Whole wafer is coated with silicon nitride. Only pads are disturb. The rest whole circuit is protected by nitride. Nitride is a very hard material, if you touch it scratch it, it does not go below the surface. So that is why it is also it stops water leakage. It also stops all kind of impurities to go through. This is a relatively thicker nitride. It does not allow anything to go through it.

So it is called excellent passivation. So, all wafers are passivated finally. The last masking is done for only pad patterns, okay. The rest whole chip circuit part is protected by nitrides. Then you are taking to whatever the package you want. First dice it into chips, put each dice

on different kinds of ceramic or metallic or whatever it is. Yes, there are two technologies there.

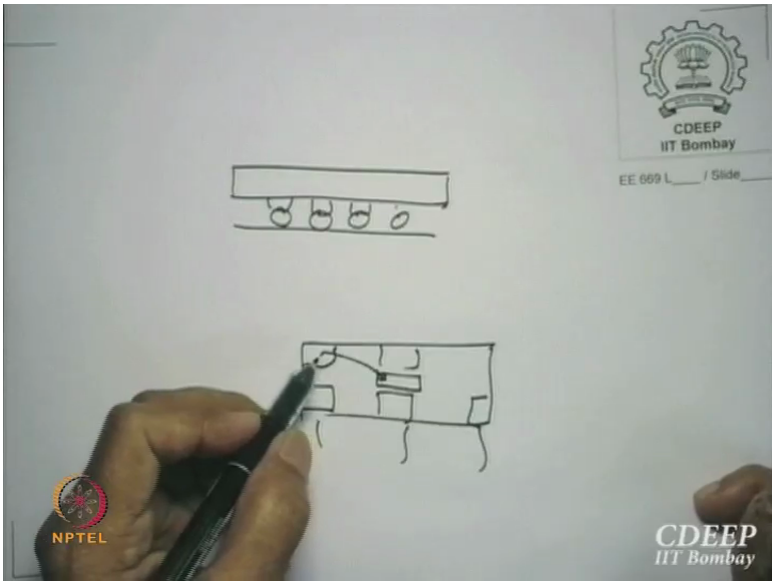
Either by wire bonding you can take from this or there is called ball bonding. You create a ball pattern and put flip chip. It is something like this. I have metallic balls and this has the layer of actual wafer so here is the one pad, here is another pad, here is another.

(Refer Slide Time: 01:15:18)



So this pad pattern is flip chipped on the ball pattern and soldered. Balls are quite thick so their inductance is very low. We are more worried about wire the way it is in a package. See this is your package patterns pads which are actually external pins and this is your chip. You have one pad, you wire bond here. Now this length is not very small.

(Refer Slide Time: 01:15:50)



So because of that there is sufficient inductance with this. Now this inductance is used in R F and analog as an inductor itself because I need an inductor. It is variable there so it is called wire inductor. But in a normal digital circuit I do not want this to happen, okay. So I do not want this connection like this. I want flip chip. I will just bond it on balls. That is how I will use it, okay. These are all back end.

But once you are passivated both sides you just put passivation. Okay, this finishes the VLSI silicon integrated circuit VLSI course as best as we could. There are many hundreds of issues which I have not touched. I wish I could have. From your side it is good I did not from but from my side I missed a few.