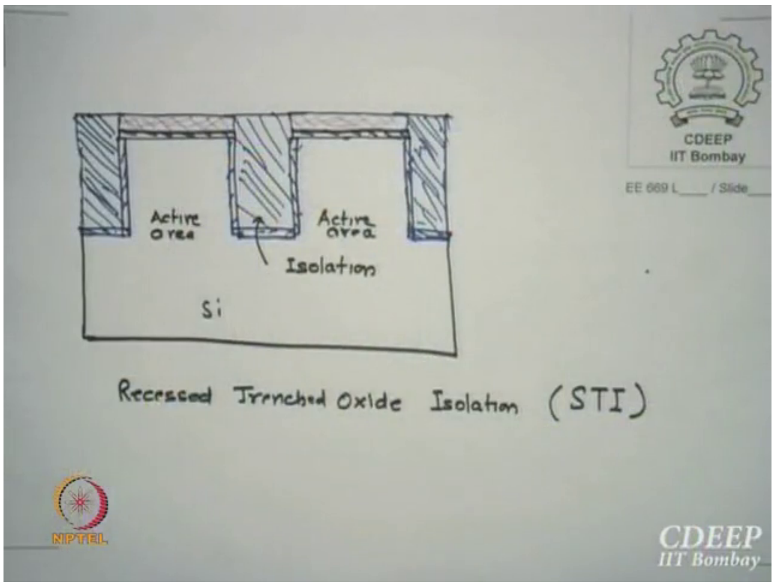
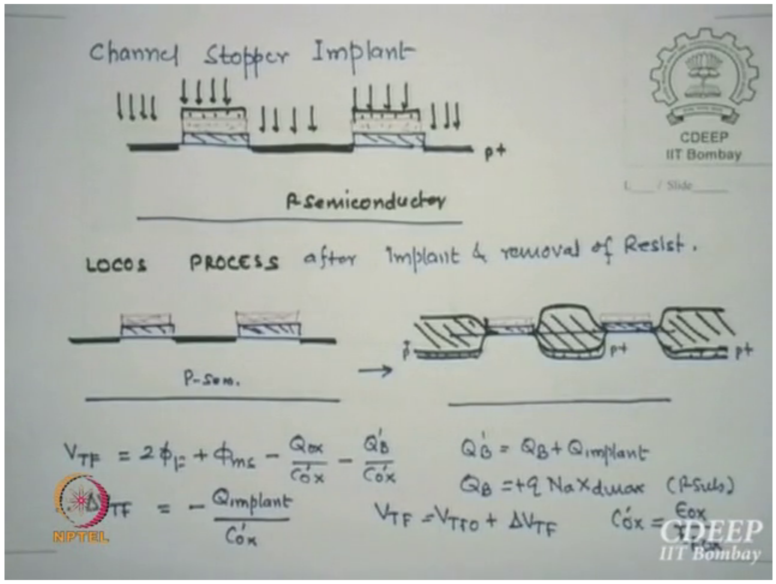
Fabrication of Silicon VLSI Circuits using the MOS technology
Professor A. N. Chandorkar
Department of Electrical Engineering
Indian Institute of Technology Bombay
Lecture no 22
Module no 01
ION Implantation Silicon IC Processing flow for CMOS Technology

(Refer Slide Time: 0:27)



Yesterday we were looking for isolation and the last we did was the Tint oxide isolation and we have now an active area which is separated by the trench which is filled up with oxide. Now we did this last line, now I will start with today if we are not doing tint and we are doing a normal low-cost process then there is another process step is needed which is called channel stopper.

(Refer Slide Time: 0:55)

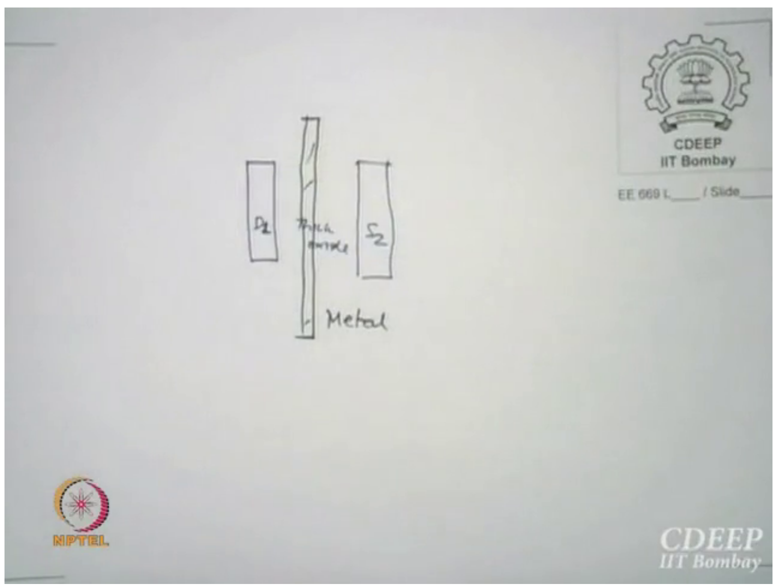


This may not be needed in STI but this is normally needed in a normal isolation oxide we grow for example, what is needed is this step is same as earlier, we have a silicon dioxide, nitride, resist and we implant now depending on whatever substrate we have in this case implanted P + I in a substrate P. This black portion is a P+ implant and this resist will stop the implant elsewhere, whenever the resist will not go through wherever the clear regions are there the implant process will make the surface below P+ then of course we remove the resist and start driving along with (())(1:55) and we start oxidation along with driving is essential in oxygen, so its oxide will be grown and since the impurities will also be driven in so just below the oxide you will have a P+ layer as you start doing the impurities will also start going down and they will be always in touch with oxide so you will have now a P+ layer between below every thick oxide or fox we have created and this is many times essential in older technologies.

In new technology of SIT probably does not need it at much, but older one has to be used with this. The reason is obvious if you have drawn I will just show you this formulation I will clear it while I am saying. Repeat, I am trying to create a P+ region below thick oxides, we have not I do not know whether the design coats have started layouts hopefully something they might have shown so in the layout of a circuit when we come for a mask, something this formula will come back to it just see if you draw the figure I will just 1st explain and then we will come back to these numbers. So I repeat, this process is necessary only if this kind of isolation is provided, if you have STI 10 then you do not need these implants.

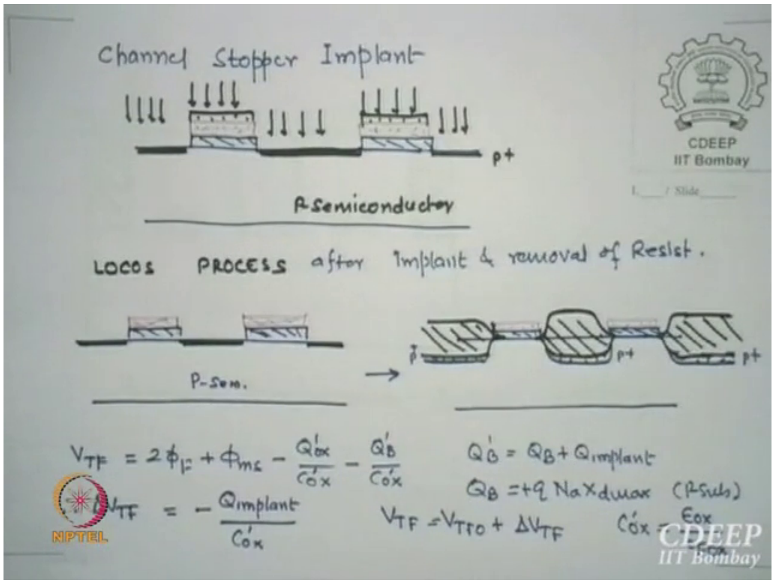
Now why this word channel stopper okay that is something I will like to explain. To some extent this question was asked in the example okay indirectly okay in the midterm is it okay? This is standard, once you know this is just get thick oxide, which we have done yesterday the only thing is now your P+ region is below all Foxes all fox regions. White is called Fox? The larger area at the wafer is essentially a thicker oxide everywhere except the place where transistors are coming therefore it is called Field oxide so let us quickly, we will come back to this sheet again, what is happening...

(Refer Slide Time: 4:35)



I do not have colour but let us take this is your gain of the transistor and this is your source of the next transistor. Essentially I am saying when I make a transistor here, we will be doing source, drain source and there will be there will be source drain source drain, so between drain 1 and S2 there is a thick oxide. You can see this, this thick oxide is sitting between the drain of 1 and the source of 1 okay so essentially this region is thick oxide everywhere and since metal has to run except the place where transistors are because it should not touch any source drain is called runner. If metal is running interconnect is running so it may run for example, in between like this it may go ahead also this is metal. Since below metal there is a thick oxide so it also (())(5:44) itself from the transistor regions okay.

(Refer Slide Time: 7:32)



However there is an issue here okay the issue starts, you understood what I said, 2 transistors separated by thick oxide and in thick oxide wherever I have thick oxide I will run the metal. So I run the metal on a thick oxide but it created some problems, If I have a metal line here okay and this was my source drain okay then I figure out whenever metal receives highest voltage V DD then this is please remember this is oxide, this is substrate concentration so this is metal oxide semiconductor okay metal oxide semiconductor so there is a capacitor formed there okay with thicker oxide. If this oxide quality is not good enough, doping is not large enough then it will look at the expression, it is 2 Phi + Phi ms – Q ox by C ox, Q ox for this may be dash I should put for field oxide + the doping of substrate okay.

However, if these 2 values are such that it does not become large enough positive value okay, how much large it should have? At least cross V DD. The threshold of the source… This is called field transistor, this transistor is called Field transistor which is built in nothing can be done on that so the threshold of the Field transistor is less than V DD H can happen depending on the oxide charges, depending on the doping you start with then inversion below this will start, the metal will invert the field oxide as well. So if there is assuming right now there is no P+ this drain through a channel field channel will connect to the source of the next transistor okay so it is shorting the 2 channels okay this shorting the 2 channels and why it can happen?

If the V TF for this thick oxide is not large enough larger than V DD then this process and typically VDD is not the only signal, what is the maximum signal one line receives if the power supply is VDD. I repeat, I have a signal running on a metal line 0 to 1 I shift, how much is the maximum voltage the line will see? It will actually see a transient depending on the inductance of the wire of the metal line, L di by dt will be the large amount of voltage will create and 2 VDD is the possibility depending upon of course Zeta as we call and it can even some time if there is a capacitive effect it can even boost further boost up. So the safest V TF is 3 times V DD but the minimum we should try is twice the V DD is that correct, minimum V TF voltage required was say operation is 2 V DD.

Now if I want to increase V T, how do I do increase the V T in this function? Phi ms increased but it is locked on so it does not increase the value of... Phi ms is 0.6, 0.65, 0.68 it will go as much as 0.04 volts maximum change, the only thing I can change his (())(9:38) charges the doping Q and AXD okay. So if I want to increase any but I do not want the rest of the places that doping should go, I only want the Field ratio to be higher so wherever field was there I put a P+ implant everywhere which will increase the field threshold of this so-called parasitic transistor larger than 2 V DD, is that got clearance.

Student: Sir.

Phaser: Yes.

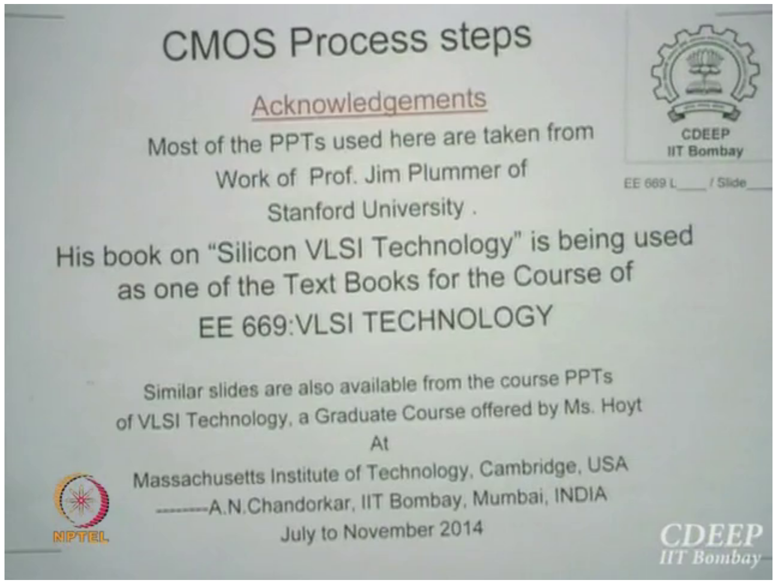
Student: (())(10:12)

Except for the transistor the field oxide is going to come everywhere so wherever field oxide is going to come below that it should be P+ okay. You can see here there are 3 regions, all 3 regions there is a P+ oxide is going to come here, here, here, wherever there is a field oxide below that there should be a channel stopper and that is why I call channel stopper what this channel stopper mean? This channel does not get connected to this channel and it stops okay because of this transistor does not turn on okay, this is essentially is called channel stopper implant. This is as I said, why I said in SIT we do not do so much because SIT is very deep relatively the source drain so it is unlikely that the channel will go all the way down and get it connected okay, not impossible but difficult.

So when say SIT is a good isolation you do not need channel stoppers that much as you need for normal this kind of isolation which was the old technique and till 90 nanometre everyone was using only, below 90 only we went for SIT okay. So this how do I get the V TF extra? You know what are bulk charges whatever is Q and AXD + whatever base you are going to put through implant, anyway (())(11:47) is pollinate area of this charge is also pollinated area is that correct, Q into (())(11:51) is Q v dash or Q implant. So I know this because I am fixing through implant how much I want, this of course without if I do not put it I said it is a V TF and if I put that it is a new V TF dash is the subtraction is without P+ with P+, this is Q implant by C ox is that correct? And it is – because it is a negative charge income so it is essentially + well you will appear.

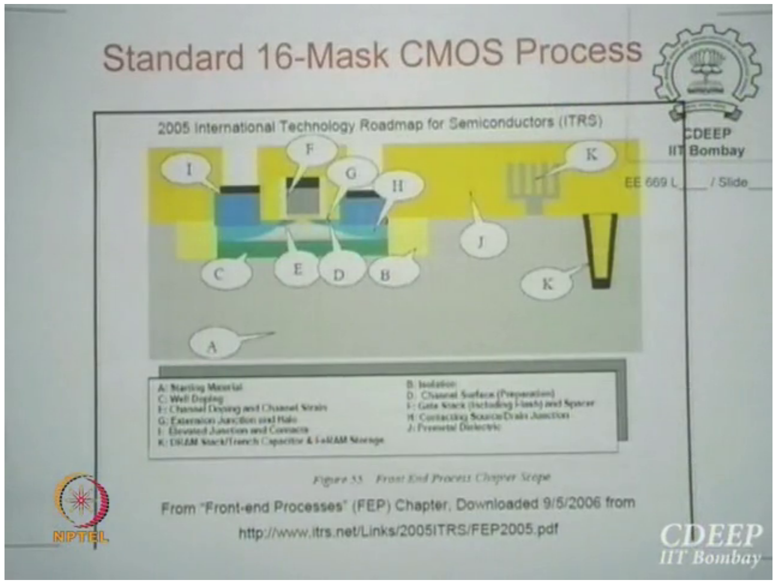
So with this value is essentially the additional threshold which you want to add to the Field transistor threshold which you would have got otherwise if there is no channel stopper. And this means that those can be now decided by how much additional V T you want from the earlier one if you add that much dose implant here, so it will give 2 V DD if you want adjust V TF whatever initial was see that it increases to 2 V TF and that 2 V DD and actually find the dose, so implant those is decided by the choice of V TF you make decision, if you want 3 V DD you may have to put higher implant okay much more dose you need okay, but at least 2 V DD should be safer okay. I repeat, in circuits particular digital circuit the transients actually are more worrisome than the... Because DC we hardly care 1 and 0 if it is there who cares only when it may transit the shall starts okay, so think of situation that in transition the circuit should still behave okay.

(Refer Slide Time: 14:06)



In the case of analog also there is issue but maybe some other day, maybe next semester course if some of you take where I teach Signals design then you see what is the... Why even there the analog part also has transient issue okay particularly if I am doing A to D converter okay. So now I have starts so far we have done it, now I will follow whatever process steps given in the book and my acknowledgement to Jim for allowing me to copy without his knowledge okay, of course similar slides are available from MIT, course was normally given by Miss Hoyt, she is also professor, all of them now have become nanosensors okay including A. N. Chandorkar and others. So these people have their old courses, you can also go ahead access to their, these are open courses so no password. Okay so I just because this is given in the book so I thought that I should inform that this is standard thing which is available in books itself or in the net itself.

(Refer Slide Time: 15:07)



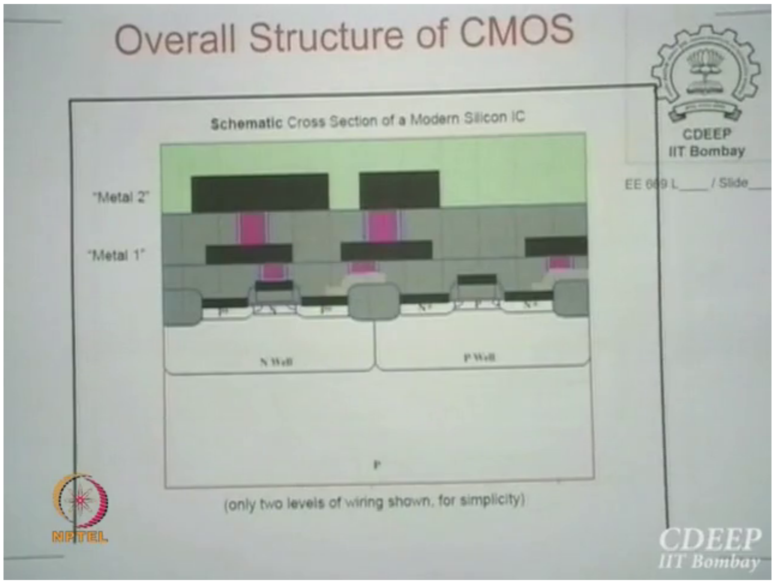
So the process step which we are going through now is called 16 Mask CMOS process and of course this is a very old slide, even in his book of 2009 he has given the same slide so I have not changed okay. You can see from here, this is one transistor should here this is your starting material then this is your isolation B, this is yellow is your isolation, we can see here is also a deep trench oxide is there but this is not for isolation, what is it? It is a dram it is a dram capacitor okay trench. Then there is a well doping C Green one we will do that now followed by channel surface preparation then we have channel doping and channel strengths silicon germanium if you add and then you have get stat and spacers as the major process step.

Please remember someone other day was asking, this spacer is not only for one reason actually spacer is a good thing and spacer is a bad thing, too much spacer thickness may worry you too much, lesser spacer may even worry more than that so spacer decision and spacer variation because on a wafer it is very difficult to maintain sidewalls. So the major (())(16:38) issue in all the sidewalls is a spacer how to create, so there are normal processes may one step we go, but nowadays we need 4 steps to actually create accurate spacers okay, maybe after this whole course last day I will show you what is the spacer technology which itself is a gain now okay. Okay so then you may contacts by opening silicon and gate and of course as I said this is a dram capacitor, this may be a resistor, this is also there is something called ferroelectric Ram FRAM so there is also a finger kind of MOS structure which is created for FRAM and of course there is a thin oxide which is your gate, which may be not necessarily silicon dioxide but can be silicon nitride or can be any other IK any other IK OK.

Okay so this is this if you are really interested (())(17:40) there is a body which is ITRS as the word says International Technology Roadmap of Semiconductors, there is a huge body some 150 people work for it, various companies academics and they predict what they actually 1st figure out where is the technology now for each kinds of devices, DRAM for all kinds and then they predict for the next year so they decide these are the problems, how it will possibly get solved and next year what is expected okay that is called roadmap. So this 2014 roadmap is already out, now you can go and look at what is 2014 people have got through and what is expected in 2015 okay, what are the bottlenecks even in 14 which they will solve probably. It does not happen next year okay they say okay we could not solve so that is the present state, in 2016 we will have the last 15th copy but that is keep doing.

So this is of course as I say standard this, ITRS is the famous web page for that, go for it only thing here 2014, otherwise it will show you 2006, 2005 any year, so for every this you ask for slash 2014, 2012 okay. But 2012 to 14 is not very much change 13-14, 15 they expect now we do not know whether yes or no. Okay so this is what I want to make, this is the ultimate what a MOS transistor I want to make 1 n-channel device, 1 P-channel device for a CMOS okay. For NMOS I only need one of them and of course there is some cases we will make resistors which will require resistor implants but right now assume there are no resistors, and capacitors are always MOS capacitors then there are junctions so I have junction capacitors so I do not need additional processing for making capacitors, but resistors I may have to okay. And also inductance then I will really be in trouble because I will need a huge area to create the coils okay.

(Refer Slide Time: 20:09)



Okay so we start with tested okay this is again the final version of what is typical CMOS will look like, this is only 2 metal layer process, the current trend is 7 metals okay so this is metal one then there is metal 2 then there is isolation metal 3, metal 4, metal 5, metal 6 and metal 7 okay, our ultimate aim is metal 12 okay but I think I do not know we will reach there are issues of thicknesses. The problem is, where should be keep VDD or where should your signal should run maximum, how much capacitance between the metal layers okay. So there are too many issues because if you increase R then whole your speed goes, you are looking 6 gigahertz from the transistor and the interconnect is actually making it 2 gigahertz so what is the point in making transistors 6 gigahertz.

There are issues so this process number of metal layers and their contacts the making the 1st layer transistor is called front end process and the 2nd onwards or at least 3rd onwards whatever metal interconnects you create that is called back end process okay and back end process actually decides the circuit or how is to be connected okay. Front end process designs the transistors, back end these are called interconnects, so major research right now is in interconnect not so much in transistors, of course FINFET is not now known for long so there is not much change has happened from FINFET of 2010 to 2014 okay. Okay so this is not a FINFET this is a normal MOS transistor, there is a P-channel device and there is an N-channel device and this is the gate oxide, this is the inversion channel in between N and P.

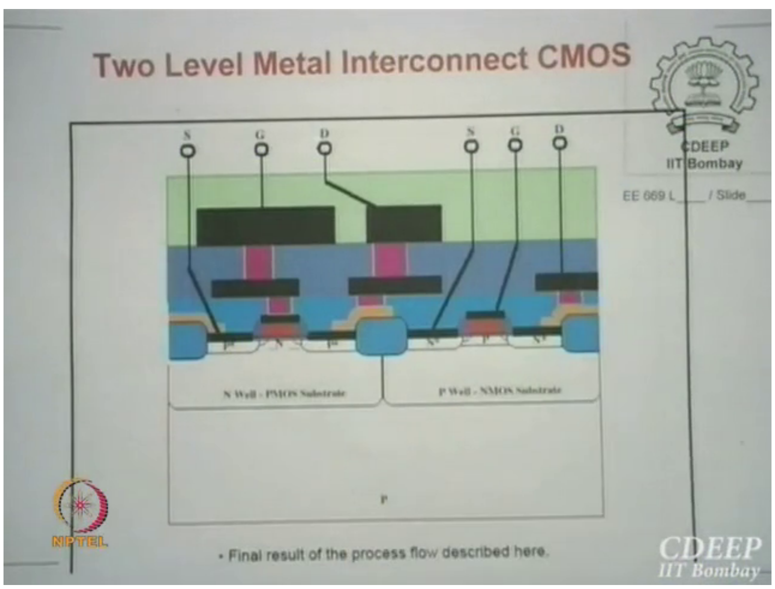
Okay please remember since I need one N-channel device and one P-channel device so I must cannot have same subset so I have created a separate region for each transistor so they are called wells. So for a P-channel device I should have N substrate so N-well, for N-channel device I need a P substrate so there is a P-well okay. This is deeper area down at least okay that is called well, well means deep enough compared to whatever you are doing at the surface compared to that at least 4 times it should be deeper okay, this is essentially adjusting the length and that decides speed. So all these gains are related to speed or all of us are now trying to see what parasitic we get which reduces my speed so all my efforts in technology is to improve speed. Is that clear, technology is varying or why everyone has to do something more, is to see that the transistor speeds are not limited by interconnects or not limited by some other side parameters okay so that is something very important.

So this is as I said a typical okay this red one let us say I want to connect metal 2 to metal 1 so you can see I have connected some red portion, this is actually tungsten okay it is called tungsten tub, tungsten has its own problems so it should be guarded by something so there is a card tungsten tub okay and can see that bluish it can be titanium nitride, we will see this process anyway. These blacks are metals so metal 1 is to be connected to metal 2 so in this oxide I must 1st create a hole yeah as the word is VIA so I must create a VIA down and then fill it up because I want the contact from the both bottom and this. Now the way it is shown here is not true because once normally wherever gate will appear that contact should not be, there should not be any contact above it so it should be always displace contact so you are here then you move then you move up and that is why areas are required because you will not put one over the other you will keep moving okay.

So you can see from here there is a contact to the drain of this transistor, this is the source this is also the drain of the P-channel device, this is P-channel device, what is this contact if this is a source, where it will go in real life? This is a P-channel device, this is N-channel device, this is the... Source and drain are identical in this particular process so let us say this is source, this is drain of a P-channel device, this is source and this is drain of N-channel device so where this will go and where this will go for a CMOS inverter? This will go to V DD and this will go to ground and they have not shown it but these 2 connections will be internal okay, I do not need to separate whenever I make this mask for metal I will have connections between 2 drains okay so I do not have to really connect taking it out.

Okay so now there is a separation has to be done between this and this, I want to make a contact please note it I have made contact somewhere here, this contact I did not push it here I went up open a window, put a stub there and went up okay so there is some mask which is to be designed every now and then that is why it is called back end process which can be decided after transistor processor is ready then we decide for a given circuit what should be back end interconnects okay. Why it is called back end because at the end device we are now deciding how to interconnect therefore it was called back end. So back end engineers jobs are not very good they work too much on software like many tools are available CAD tools but they are the ones who make money.

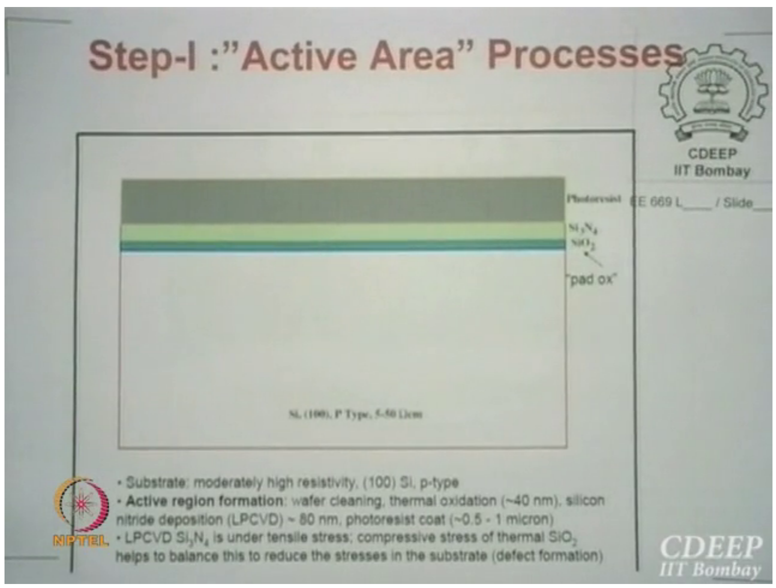
(Refer Slide Time: 26:36)



Okay WAHI CHEEZ HAI USKO THODASA ABHI connection DIKHAYA HAI UNHONE. This is the source of P-channel, this is the drain of N-channel or maybe source, this is the source, this is the gate and one can take this contact up up, this can get up, this can get up, on the surface all contacts can be made thin as pads okay so we can create pad connections. So this is the final device we want to make and we have already done so far what? Isolation, we have not done anything ahead, we have only made active area where transistors are so whatever we have done is, we have opened these areas so far okay, this is our isolation oxide fox and as I said there may be a P+ below this, N+ below this.

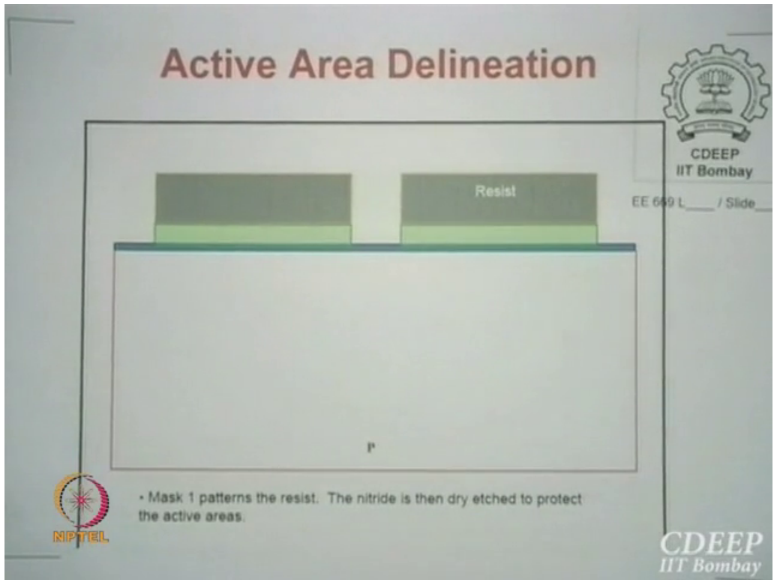
So now we have 2 need mass because for this P-channel device what implant I should do, Channel stopper should have N+, for N-channel device I should have P+ so another mask, what I am doing here I should not here, what I am doing here I should not do it here so another mask, so Channel stopper is not mask less, you need 2 mask to do that okay so there is also a catch in the whole system so that is why we have not shown it, he is assuming this doping is sufficient to take care but it is not true.

(Refer Slide Time: 28:11)



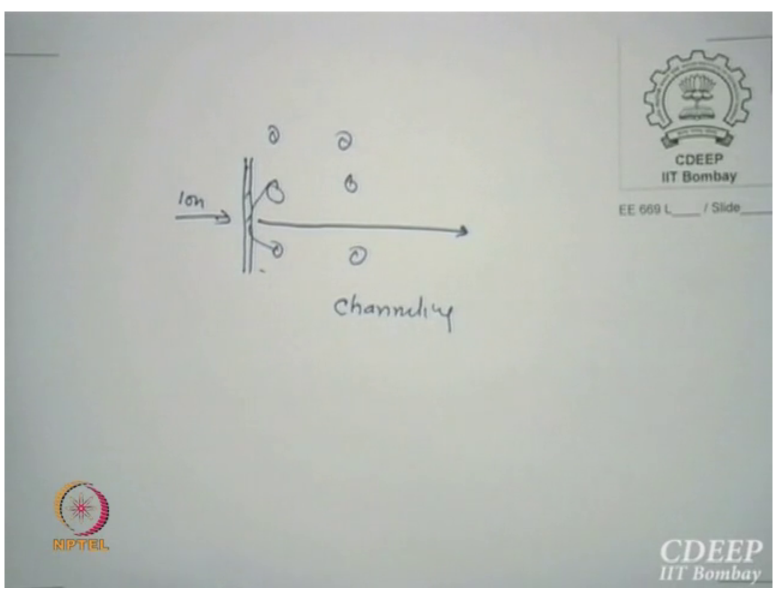
Okay so this we have done so I will just repeat, yesterday we started with oxide, nitride, photo resist, 1st step yesterday, this is I am just repeating what I had actually drawn and shown you yesterday. This is 1 0 0 wafer okay, silicon nitride was deposited as I say by LPCVD and I do not know how much he is talking of 80 nanometres but it is okay 800 armstrong I said, I am still comfortable with Armstrong than with Nano so just 10 times that; 80 nanometres is 800 Armstrong, this is low-pressure CVD we will look into this, this 1st oxide is thermally grown.

(Refer Slide Time: 28:58)



Then we have a 1st mask and there was a pattern shown to you in which this was retained and the rest was etched out okay. Once of course this oxide can be removed or can be held, this was held because if I am doing Channel stopper I need something. Oh forgot that day I am sorry maybe quickly 1 minute I will come back to implants and you just you have drawn it yesterday. Sometime which I forgot in hurry but I think that is very important.

(Refer Slide Time: 29:37)

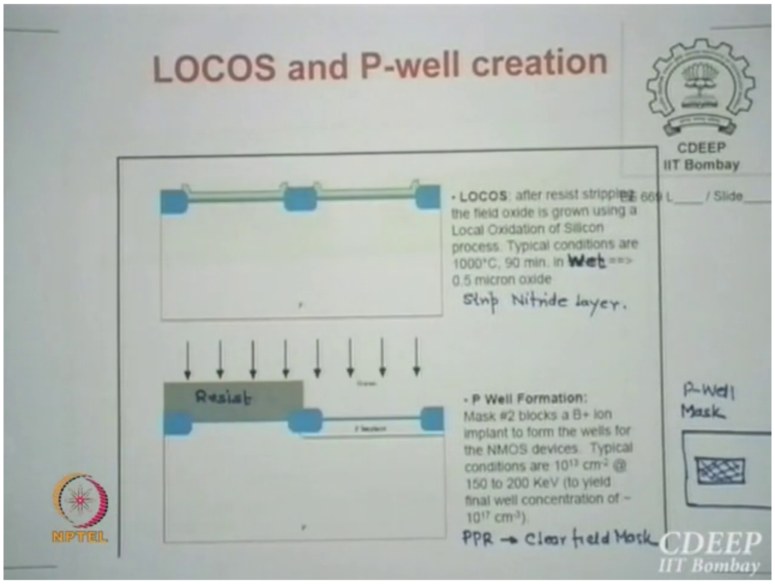


One thing you have to understand, when I am doing an implant and there are number of atoms, you have silicon there is a possibility that if the ions which you are passing like boron which is smaller actually may not interact with anyone okay and just go and go okay till it somehow find some dislocation or something where you hits okay this is called channelling is that correct, this is called channelling so extra range is seen because atoms go in between like a corridor, it just goes in, no interaction. Now this channelling have an issue, if I put a very thin oxide her, oxide is not a crystalline material so what will happen, the ions will hit and will when they come out they will not come into straight ones so they will actually get in different angles and therefore more probable to form a (())(30:41) profile, is that clear to you.

Is that why this oxide is retained? Because if D Channel always that incoming ions is that clear, this thin oxide is D generalises the incoming ions, oxide is very thin so most energy it can pass through the crystal D-channel itself okay so this is why the word which I was the figure which I was showing you, is that okay what is channelling? It is a corridor of atoms periodic you said it is periodic if the gap is 1.62 Armstrong ion is smaller it can go straight so that some ions may reach much deeper compared to the others okay so they may actually change the projected range okay. So to do this normally thin oxides are retained okay, someone asked can you... yeah I can etch I can always etch it okay, when I was etching all others I can etch that as well okay.

Then someone also asked that when you etch nitride, the etchant is same for oxide, etchant is same for that so the way we do ionic etching or what we called as anisotropic etching so I can etch what we call etch top, as soon as the oxide starts the etchant goes down okay, energies are so adjusted, only nitride is etched out okay. So there are methods in which oxides can be... In case there is no oxide, before you have 2^{nd} mask here you have to remove resist put oxide and put again resist so it has done masking but that is costly okay.

(Refer Slide Time: 32:38)



Okay so we are now ready for active area then we did LOCOS as yesterday we said and then of course we strip the nitride layer after making LOCOS and then we start the 1st process which yesterday after this LOCOS has been made you remove the resist this photo resist as your nitride so you have now a clean silicon surface okay with isolating islands okay these are active regions, the rest is thick oxide everywhere. Now I use the 1st mask, have already shown you need a P well to create N-channel device okay so I said fine I actually put a resist on this portion using this mask, if I am using a PPR what mask I am using? I want to retain these areas so light should not pass which means dark area this area should be dark in the mask and the rest should be clear so this is a clear and mask with a window of dark window which will retain these areas okay.

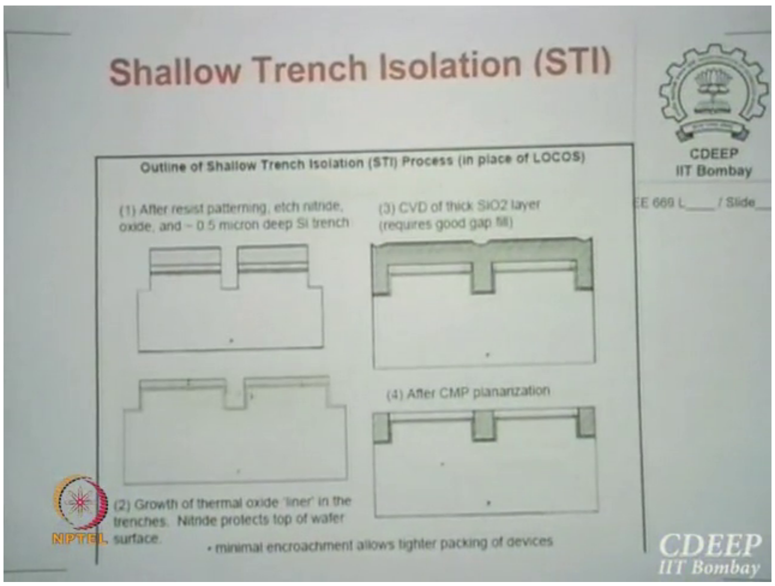
So I retain it since that is a resist it is excellent mask for any impurities so boron is implanted down and it will P implant below okay P implant below. Now the typical dose which is required is 10 to the power 15 per centimetre square and the energy is 150-200 kV, it will typically yielded to a final concentration of 10 to the power 17 okay P wells... 10 to the power 50. If you use NPR that should be clear window, window should be clear and the rest should be dark there is [no audio] that decides the dose. Now this energy is also decided by the (())(34:31) range you are looking for okay, please remember which one should have a higher energy phosphorus or arsenic or boron, N or P? Which are the larger range at same energy?

Student: Boron.

Student: Boron.

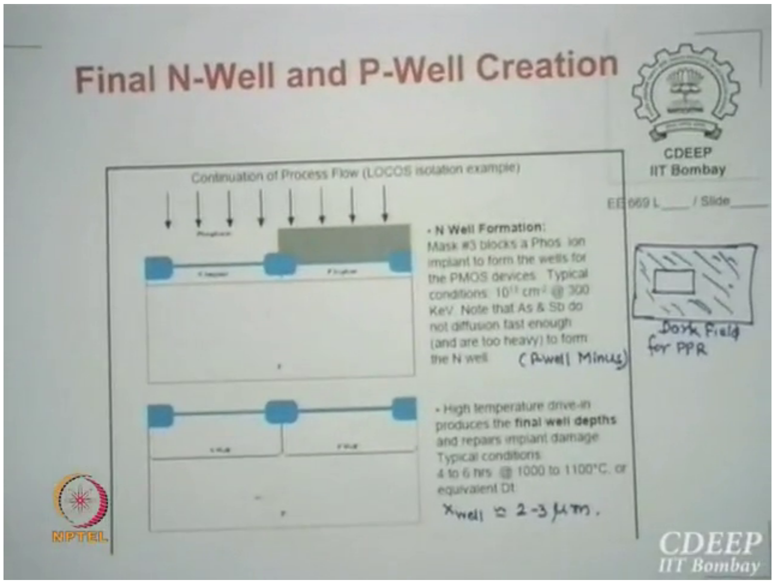
Boron, it is lighter ion it will go deeper anyway okay, whereas phosphorus so if I implant the same depth then what should I do? The implant energy for arsenic or phosphorus will be always higher then boron so this they show you they do not have but this is done is simply because equivalent value they find out so that same well thickness are available okay. Okay so what is the 1st thing we did? We created area active area and in that we make a P-well okay implant, right now it is not a well it is only implant them.

(Refer Slide Time: 35:33)



Yesterday of course we assume thisSTI this is given in this so I just printed but do not use that we have already seen that.

(Refer Slide Time: 35:41)



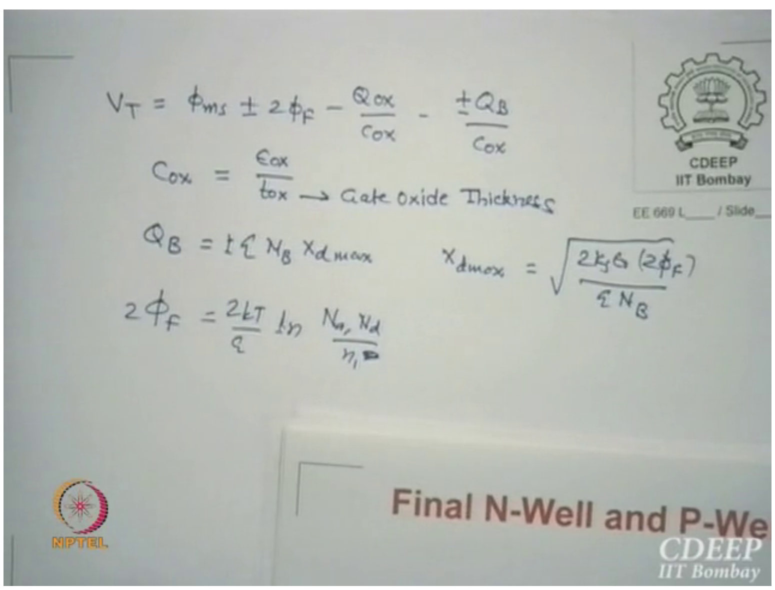
Then we do a mask to make a N-well area okay so what should do there? There is an interesting name we give the 1st mask was called P-well mask, for creating N-well we actually have a mask which is called P-well – mask okay P-well – mask so wherever I want implant earlier I do not want implant now and keep outside anyway will block everywhere so I am not worried about thick oxide area. So I have just taken a copy of that, is that clear? I use the opposite resist, she is asking is it okay, I just take the opposite of that complimentary mask as they called and just use the other resist so I can open this okay or rather same resist okay same resist, if I use complimentary mask then I do not need resist or you change the resist with the same mask also the either way.

So now this area is open and this area is blocked, now do arsenic or phosphorus implant, normally phosphorus implants are done for well, why? They can go deeper, arsenic is a shallow implanter or it has no diffusivity so it will not go too deep, dt products are much smaller so you must put deeper if you want you, put phosphorus. Okay so this is also similar dose okay typically 10 to the power 13 and look around I just (())(37:17) the phosphorus implant is done at 300KV, where was boron done? 150 to 200 so it roughly 2 times or 1.5 times the earlier one we do so that they come almost same time okay.

So I am done I may start (())(37:36) in oxygen to some extent or at least in nitrogen for a long typical this is you want depending on what is the time taken and temperature chosen by what method what is the reason I am using some cycle? XJ, I want dt products so that I get so much XG, I have 2 root RP square + + 2 dt so I just my dt product so that XJ is what I want XJ is what I want okay. So typical values are shown here which is around 2-3 microns or junctions but any value it can be lesser or larger depending on the technology node 1 is working at, the current technology node may be less than micron 8000 Armstrong but 11 nanometre it will be even lower so we do not know what is it going to do.

So after P-well, N-well (())(38:34) what is the advantage of this? It also (())(38:39) all damages activates all impurities and drives them in okay so this is a major step of creating P-well and N-well in one go okay. Okay so we have now active areas in which your area of P-well, area of N-well where transistors are going to appear. In the P-well I will get N-channel device and in N-well I will get P-channel device okay. Now the next problem is we have been doing this often expression okay this expression can be... maybe I can rewrite this expression.

(Refer Slide Time: 39:31)

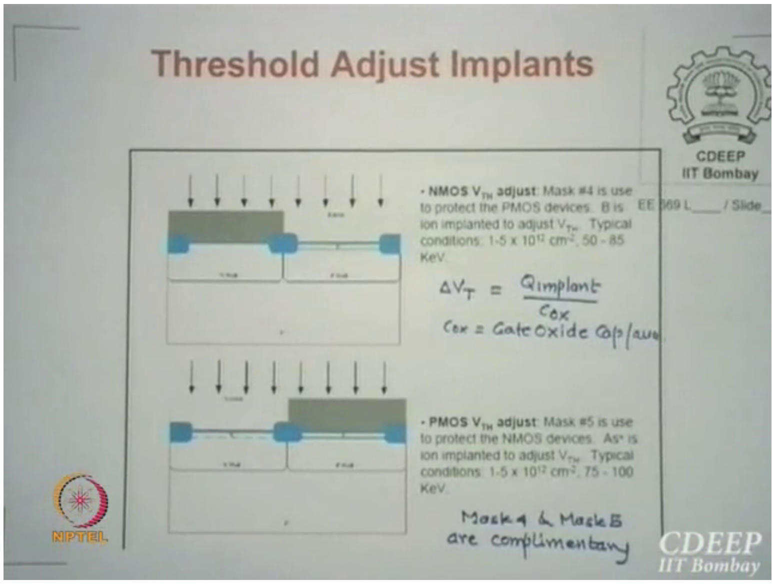


We also know for any transistor for a MOS transistor, VT for P-channel device or N channel device is Phi ms + – 2 Phi F – Q ox by C ox – + – Q B by C ox and C ox is Epsilon ox by t ox where t ox is now gate oxide. Q B is q N B, N B means either N a or N d, X d max, X d max is under root of twice K s Epsilon 0 2 Phi F upon q N B. So I know everything, Q ox is of course process dependent, Phi ms is also the metal which I use which I know anyway Phi ms, 2 Phi F is 2 kT by Q L n N a or N d by n i. So I can evaluate the Fermi energy Fermi potential, I know Q ox, I know C ox and depending on the doping I can decide my threshold okay. So depending on this so normally I have a P-well how much doping I was telling that, 10 to the power 17 but this may not be required or this may be smaller or larger depending on the node you are looking for VT requirements.

So I must now put something additionally in the Channel area which is essentially required for even VT value, is that well clear to you? The N-well, P- well dopings are not good enough to control the threshold of N-channel and P-channel devices so I will now use lower implant of exact doping of my choice or in P-well as well as in N-well, which will decide my threshold which is called threshold adjust implant, what is it called? Threshold adjust implant, is that clear to you that why we still, we are not go banking on the P-well, N- well because they are driven in, their profiles are now very flat one so I cannot decide and the channel region how much is accurately uniform doping there. So I actually want to know what is VT which is decided by whom? By the technology you are told to derive and we said this much current V GS – VT, now I have design all my circuit based on this current or so much (())(42:38)

Now you suddenly say no no, V TH is 10 percent different or 20 percent different then my whole circuit may perform or may not at all okay, so my worries are that I must get correct VT of course still it will not be correct but closer to correctness, nothing is absolute in the process everything is statistical.

(Refer Slide Time: 43:11)



Okay so after the N-wells or P-wells are done I now go for threshold adjust implants let us say 1st I will do it for N-channel device so what area I should block, all being well areas or sorry all N-well areas where P device is to come should be blocked and please remember this is always show you only this so the other portions are open it does not matter because there will be what there elsewhere, thick oxide so it does not really matter if I mask it or I do not mask it is that clear? But this window must open because there is where implant is going to be perform, is that clear? The rest area is anyway field oxide so implant does not go through what is that thickness we have to adjust?

For the highest energy implanter how much is blocking you want may be 8-9 or 10-9 okay so adjust your thickness of field oxide such that nothing goes through it in the below surface, is that clear that is why we other day showed you how to calculate how much blocking it can give a layer can give okay. Okay so now I open a P-well area and I want to adjust nothing will go below that that implant was done prior to oxidation did you get the point? Impurities what you have already put there cannot be changed now is that point clear, below oxide whatever you have put there is nothing can come out now or in oxide ions are not ions actually they will discharge through their bonding so there will be randomness in that and they will not contribute and if they are away from the surface they will not contribute to VT.

Fixed charges are only when it attach silicon or it is a dangling one, these ions are much clearer than available SiO bonds so in general nothing will pass through this and they will lose the charge and sit in a interstitial side not contributing to anything is that clear because that is an amorphous material silicon dioxide is not crystalline is that correct so there is nothing much except the fixed charges which occurs close to the interface they are not inside the interface okay. Mobile ions move because you are charging them and you are asking them to move, otherwise ions will lose the charge as you put in the wafer at the end so they will sit from there without contribution to anything okay.

Okay now I do a threshold adjust for (())(45:44) devices boron is okay now you must remember since it is a P-well, the implant is also P-type why? Because it is an N-channel the threshold is adjusted by P-type doping so I do a boron implant in the well okay somewhere in the well. And essentially whatever is Q implant I do divided by C ox is typically the VT I can adjust at that point okay, is that clear. You have initially P-well so I know what is VT with P-well do dopings then I do implants and I know how much is additional VT I have got through which is my implant VT is that clear? It is called shift, I am actually shifting the VT okay from the P-well VT to channel VT okay so I adjust my VT of the channels okay.

So firstly I did it for what? For a in the P-well means I am doing N-channel this so did boron, I will use complementary mask or complementary resist the same mask either way okay. You keep mask change mask then resist change, complimentary mask resist is same so I open this area and now what I implant? Arsenic, this is very important because now I do not want deeper things to go so I want very very shallow implants. 1^{st} actually this I did not tell you, every process we go through there is test wafers are there on every process so before we start the next process we actually measure the last process what has happened. So I have vast capacitors of every time and I have junction device from every time, resistors of every time, there is a test area on mask I keep monitoring test areas.

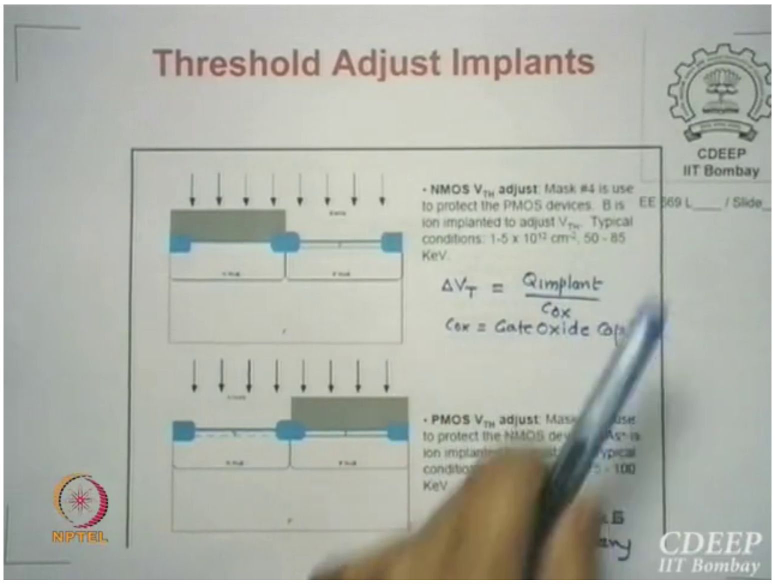
I do not use that again the same test area in the... I use the fresh, so out of say 8 times I want testing so there will be 8 test wafers, only one at a time I will take and only that process I will measure and then I will not add to that and next time I will use the 2^{nd} one, is that clear to you? So this testing is done every process time so I know VT is how much was with the P-well, I have a mask capacitor for that okay on my test so I will monitor that and I have figured out and this is not 1 test area, actually on a wafer each should have at least 5 test areas and on a wafer you can say if 100 10 by 10 area or 25 by 20 so many test chips are available. Test chips are kept in the corners and wafer top bottom base centre so 5 so there are too many test areas available to you for testing okay/

Even the final test can we do the whole circuit is a (())(48:53) on a test like I use a flip-flop many times so I will have each kind of flip flop available on the test area so I will actually monitor before circuit is tested because it is too costly to test full circuit so I will test every block. So our process step keep having test areas and if you have forced something you add additional wafer to do some more test okay only one process step you want you add more extra steps extra process wafers. Some wafers are kept constant till the end that is why the number of test wafers could be 8 to 20 depends on how many times you want to test okay. You can keep adding also additionally in between okay because now last step I have finished and now I want fresh from here, the next process what changes so I can only add one more wafer with a test area okay.

So I keep adding test area, I keep test wafers and also I keep monitoring the combined once and change once so I know what has happened to the last step to now okay, so this is always tested and only then... But this is normally 1st done by a CAD tool means you are doing a supreme Central kind of thing and you know roughly what is back to happen okay but I do not trust that result, I will actually test physically everything okay. Or you said you do spend so much resistance, I will monitor so much area sheet resistances whichever I am creating. You say doping I say I will measure (())(50:28) and find how much was actual doping gone there okay, so there is test area on every chip and every wafer lot numbers and so many wafers are kept adding so testing is during processing is very crucial for success. It has done once do it and hope for the best nothing happens okay then nothing will happen okay nothing will happen no circuit will be seen.

So this is very crucial in our poll process line that we keep monitoring what is happening okay and we keep comparing it with what our central result and we say okay so I actually modify central result now and if I monitor something many times and if I find this is the value I am not getting so I will actually change the software there I mean values there and then the new values I will get from them and next process I will modify that. It says online modification can go of course one standardise even test is done but central has not used then they know this will work but (())(51:34) takes lot of time, is that clear? So process standardisation is a gain and but money is only on if your process is standardised okay then your 200 wafers will all are well.

(Refer Slide Time: 52:32)

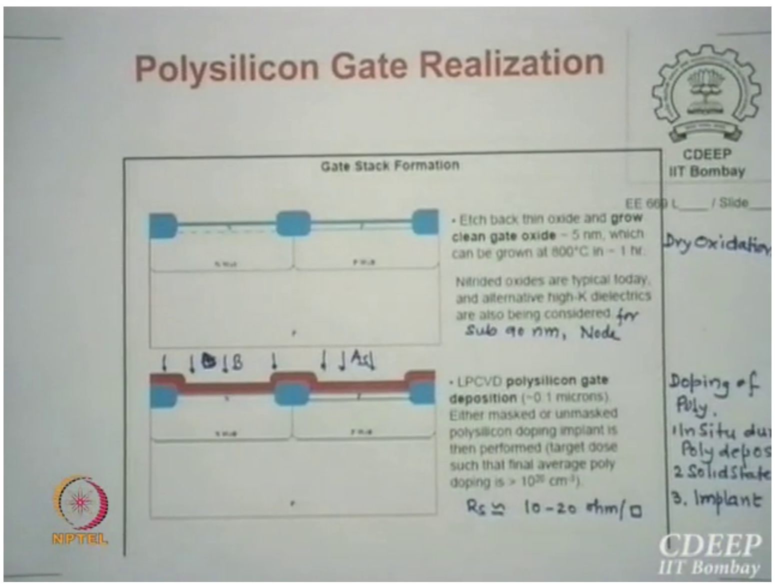


Okay so similar thing I did one for N-channel device is similarly, I did for P-channel devices this is called threshold adjust implant, please remember mask this and mask Phi R complimentary mask or complementary resist whichever way same mask with different resist or say different mask with same resist you can use it either way. So they are sometimes called – mask, P + – mask, P + HAI FIR – KYA, USKA complimentary HAI. You can also see here in the case of boron how much was the energy? 50 to 85 KV, how much is for the arsenic? 75 to 100 is that clear, so this energy adjustment is for similar that is in both cases is that clear? I want same both sides okay.

So energy I know energy is proportional to the duct range so please this take from me that this energies are (())(53:01) are adjusted to make equivalence of both devices many a times okay so that they are, I do not want VTN should be different from VTP okay, there is no harm in actually having them different but if I am making a CMOS circuit I will not like the high and low values with different okay so I want transition should be similar for both P-channel and N channel when it goes from higher charging and when it discharges I want same process to happen so I will always see to it that VTN's are same as VTP,s but the size I cannot do so I know that V is double or 2 and a half times so I will actually size double the P-channel devices or 2 and a half times just to maintain similarity of P device with N device, is that clear?

So this equality of circuit equivalence is always use the technology because we want equal everything from P or N. Okay so so far we have done we are still to do a transistor we are still only on first implant creation.

(Refer Slide Time: 54:19)



Then the okay then this thin oxide was removed because why we kept that thin oxide? Channelling P channelling okay now you remove that okay and then I grow the most important process of whole MOS transistor is creating a gate oxide, of course instead of gate oxide it can be any HIGH k or it could be nitride. What is the advantage of nitride over silicon dioxide?

Student: (())(54:44)

Professor: No no no no no no.

Gate oxide MAI gate replace KAR RAHA HU... high K, it has almost double that of silicon dioxide okay 7 so dielectric constant is double that of silicon... are not double, 3.9 and 7 but roughly doubles. Whereas, high K means really I want 16, 22, 28, so I may go for (())(55:15) oxide, zirconium oxide and may go for a lanthanum oxide, I can go caladium oxide, I may try many mixtures tantalum oxide, I have many possibilities which I can try okay. Okay so I agree on thin oxide and on that immediately okay this of course is an optional state that I have shown you, what kind of oxide it should be? Good oxide is which oxide? Dry oxide, dry oxide is the best oxide okay. But high K there is no dioxide that is why I will put whatever it comes, if you are doing SiO_2 based, you always go a gate oxide with dry oxidation.

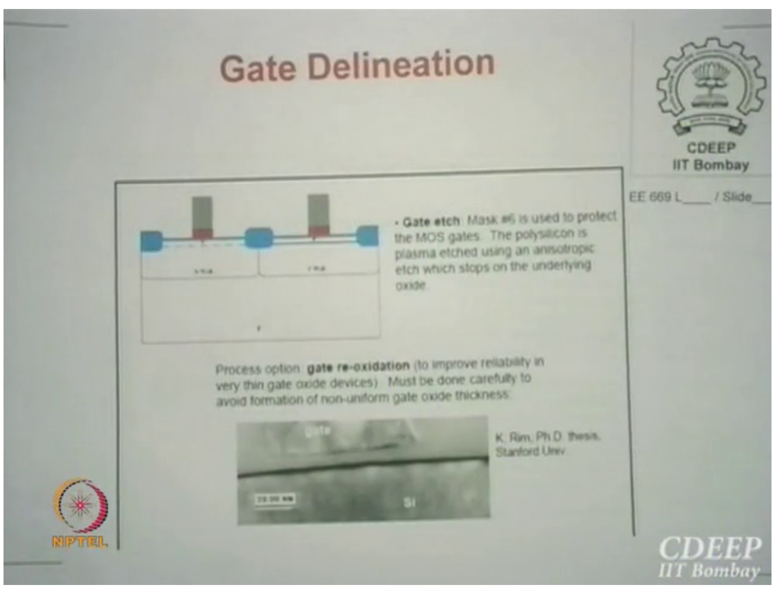
2nd reason; as new technologies came the thickness of oxide will also come down okay from... Earlier we had 400 Armstrong gate oxide on 5 micron process then we went as you started going 0.25 you want 14 to 17 Armstrong of oxide and finally what has happened now that they want less than 5 Armstrong of gate oxide then only we look for high K. So I cannot have one monolayer of silicon dioxide so silicon or oxygen either so I cannot create so then I look for that is capacitance value say okay Epsilon I can change T I can change ratio okay that is why high K appeared. otherwise nothing better than silicon dioxide, it is ideal into B for everything okay. Okay so ABHI then after this gate oxide I deposited using process called low-pressure CVD LPCVD, we will look into this process later and polysilicon is deposited.

What is the reaction of poly will be? Poly is normally assumed to how do we create poly? Silane, so you have a Silane which reacts which actually removes the hydrogen out of it yes then what should I do? I have SiH_4 and I want hydrogen to go away. HydrogenKAISE ATA HAI, ISKO dissociate KAISE KAR SAKTE HAI? Hit that is cracking the Silane means you hit the Silane and why then breaks because according to Gibbs free energy the energy of formation becomes less than energy of... T Delta S becomes higher so it dissociates okay, is that point clear, so simple as that you following in making poly silicon depositions. Some people do 2 mask here actually doped the Poly right there, is that clear?

This poly itself can be doped by 2 mask by 1 this side mask, the other this side mask so I can make P+ poly, N+ poly but many times this is not advised okay because anyway during the gate delineation I am going to implant for source drain so I will prefer to use that rather than this but there are processes books we will see where this step may be there so do not ask me why they have done it, they may have done it because they have pre-doped the... And that is very easy because if I want to dope a poly during this poly deposition I pass the phosphine gas for making phosphorus dope, arsine gas for arsenic doped and diborane for boron dope. So deposition of boron gas is much easier I mean more boron silicon by just CVD okay but I will prefer not to do this as I am saying but not that every company follows me so I am not saying, some company do this as well okay.

So I just showed you that he has also say either mask or unmask polysilane doping implant is performed. Now this is something as I said gate added, the only thing we worry that whenever I do implant on poly I want a very large concentration of poly, why I want to be heavier dope? It is a metal replacement okay so I should be as conducting as possible, the best of whatever people say 10 to 20 ohms per square is the only sheet resistance one can attain never below one or 2 which is metal can give 10 to the power -3 to -5 per square, this cannot give below 10 okay. So all (())(60:04) done it cannot replace metal, so we will do some mischief, we will and some metal to it to form a silicide, which will have better sheath resistance okay.

(Refer Slide Time: 60:27)



Okay as I said this process right now I assume only poly has been deposited some people may dope, some people may not okay. So far so good poly has been deposited on gate oxide but there are no transistors so first time we make another mask which is mask 6 to make actual gates. But what is essentially gate? Please remember gate is up to wear poly every poly region was gate oxide, every poly region would have gate oxide below so I somehow used the 6 mask with lithography as shown here this is the resist and I etch poly from everywhere else okay but I do not etch thin oxide. Even if I etch I do not mind but for an implant I will prefer to hold that okay, is that clear? De-channelization is necessary so I will keep this thin oxide.

So I want a vertically exactly etching kind so normally all liquid etchants or wet etchants are isotropic in nature, they attack all sides I want vertically down so I do anisotropic (())(61:44) which is a essentially what we call as dry etching okay or ion etching. So I create gate, what is this? The length of this is what is what it will be for the transistor? Channel length okay and the 2nd dimension which is not shown here in the inside what will be that, the width of the transistor, circuit people are only interested in W (())(62:13) so for them if I achieve this thank you very much. Of course there is also VT, they are looking only for V no they are looking for many things, they look for mobility okay what is mobility here at the surface that is very important to them Mu C ox.

So all of technology is transferring to device performance and since device is going to give me circuit performance, technology is connected to circuit performance which is our ultimate aim, this whole processing is not ultimate aim this is all done for making a circuit go okay. So ULTA MAT SOCHNA KE YEH KYA HAI thing I have opened like this, there was a time when designers used to say I want this and technology is to say well this much I cannot do for you okay, I cannot give this mobility I cannot adjust threshold so much I cannot give anything of this kind so then the designers used to crib oh you know you are not able to give me what I want. I want 1 million transistors, he said no I cannot give more than 100 1000 transistors. The things have changed around 2005 onwards, now we have number of transistors available of uniformity in billions, find the system which needs that okay.

So the question is now opposite, how much designers can think system designers then circuit designers so 1st system HI NAHI SOCHENGE TOH device KYA SOCHENGE, then what about crib anything oh YEH process AYA KAHA SE. So why are we improving process is still unknown to me, I mean this why Intel wants to go for 7 nanometre is a fun because they are thinking that they will use their microprocessor designed to much higher depth but who needs a let us say 10 gigahertz microprocessor or what I do not know except for a videogame, other than that you do not need anything okay.

Otherwise in real life you do not need 7 nanometre process anyway no circuit which you will build or system you will build, 45 nanometre is majority of circuits can go, 80 percent of circuits will require no more than 45 nanometre, some 20 percent may require 16 nanometre or 22 nanometre or maybe 1 percent will require 7 nanometres, so are this invention going on for processes only for the sake of that 1 percent, my thinking is like this okay this may not be agreed by companies but you can see I am right because most of the companies boundary lines are close because that is what they keep thinking OH HOJAYEGA HOTA KUCH NAHI HAI.

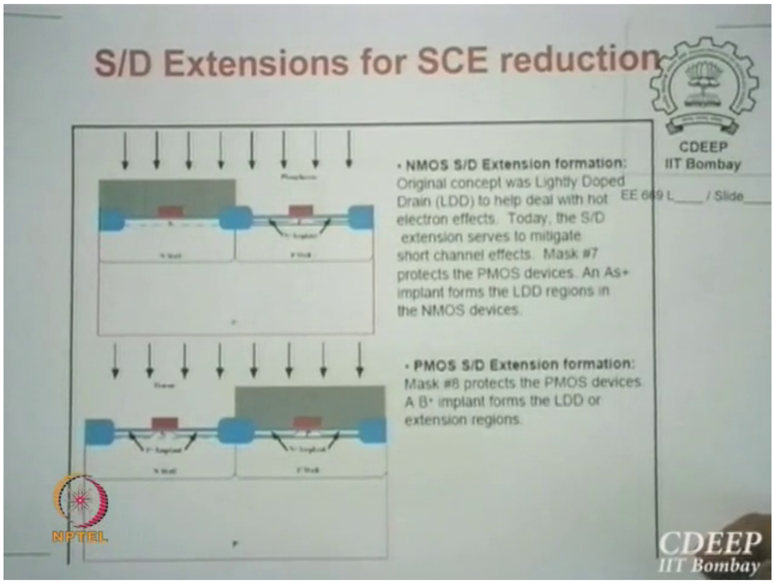
So right now my course of course... The course person I will keep saying technology is the only thing but in real life it is the system designers which actually should play better role because they are the ones who is your... Okay I want so many functions to be done in so much time frame then only I will require to do something nothing is happening. So please when you do the next courses you pay enough attention there because they are the ones who should drive us, technology KYA HUM APKO KUCH BHI DENGE anything you want I can do now but their problem comes oh I do not know what to use with, (())(65:52) there is to learn cash will require say 8K then they say no I want 64K okay 64K DIYA no no no no 1Mb okay 1Mb DIYA, you do not know how many exchanges you need on a processor more than 1Mb, you do not have even architecture which are required that much exchanges.

So just L2 BADHANE SE FAYDA KYA HAI, BAHAR DIWAN PE ANA HI HAI APKO HAR DISK PE ANA HI HAI TOH FIR USKO BADHANE KA KYA KITNA EXCHANGE KARNA, KITNA RESISTOR DAALNA CHAHTA, just because you have resistor you keep adding resistor. Aisa actual architecture people have not come out with anything great so far and (()) (66:35) JO BOLA WAHI WE ARE COMING, so something different unless happens do not blame technology, technology has reached its peak and nothing is stopping them okay. Maybe I will I do not know whether time, what is this LED business has come? 90 MEIN BANAYA GAYA THA, AB KYU importance AGAYA USKA.

Okay so now once I delineate the gate, gate delineation means sizing W by L for the gate is called delineation of the gate okay so I have delineated the gate. Once I delineated I remove the resist, of course YEH actual figure HAI Stanford University KE KOI Ph.D Thesis se Plummer

KE student KI okay you can see this is the gate he has shown okay, this is the ECM picture okay is that okay. Now there are few things more or important in 16 mask process.

(Refer Slide Time: 67:39)



Now we use another mask 7 actually to do what we call source drain extensions and this word is very important. This was required because we said please take it wireless local loop always create noise so now we want that many of the MOS transistors had when we did for scaling down to a lower technology nodes, we figured out that the channel length became so small and the voltage did not scale okay, there is no voltage scaling going on same as channel length scaling so the field across the source-drain and across the oxide has increased because of small one-to-one scale okay. If it is one-to-one scale I do not need anything, I will just reduce everything point it should perform that is what we earlier did but now it cannot be scaled down.

So we figured out that if your channel length is too small and the length being smaller and voltage is maximum the electric field even lateral as well as vertical is very high okay. Where is the maximum field in the channel? Maximum potentially that the drain so the maximum field occurs at drain so carriers starting from source get accelerated to its maximum velocity generally saturated velocity to nearby drain. And since they are the most energetic carriers and their oxide is thin enough the electric field is large, some of the electrons or carriers either carriers can travel through this oxide on the gate okay this is short channel effect which is very very dominant effect, is that clear to you?

This is very very dominant effect is high energy tunnelling going through because the barrier is typically of silicon dioxide is known and if your energy is high enough it can climb the barrier or it can go through the oxide if the oxide is thin by tunnelling, this process is called thermionic field emission okay. So TFE can add to your carriers going into this, what is the problem carriers going to oxide? They will shift the VT charge because they are sitting right there okay if VT shifts then everything shifts so I am worried that tunnelling happens. There is a model which one of my students created is called Lucky electron model, how do we decide that how many carriers will go? So there is a probability theory which we created then which said that okay this is likelihood to this.

This method is used where in ROM, Flash ROMs, E square PROMS or EPROMs many new this is exactly what we are trying, we were actually trying to push and we do not want them to come back so we put another gate and no connections so LATAK JAO UDHAR so FOTH KAR DIYA USKO so that is a flash ROM, flash of course to erase but ROMs EPROMS. Okay so I now do some kind of I figure out at the drain end and since transistors are normally symmetry, why they are called symmetric? Source and drain are normally interchangeable, in some structures we now insist it not be there but otherwise they are normally interchangeable so whatever source for this device, if I this will be source this will be drain so there is nothing great about it.

So you find at the edges I must have lighter doping, why I should have lighter doping because if a normal P-N junction if you see, the fields are if the doping is smaller, depletion width spreads you no depletion width is inversely proportional to 1 upon N so if N is smaller than the depletion width spreads so that electric field also spreads so near the surface electric field is not very maximum okay, it actually go within the depletion there, this is therefore you need drain region closer to drain area should be lightly doped than the drain itself okay that is called LDD Lightly doped drain. So this process is LDD creation okay, what is LDD? Lightly doped drain and why do we need it? I want to avoid short channel effects, yeah I want to avoid short channel effects.

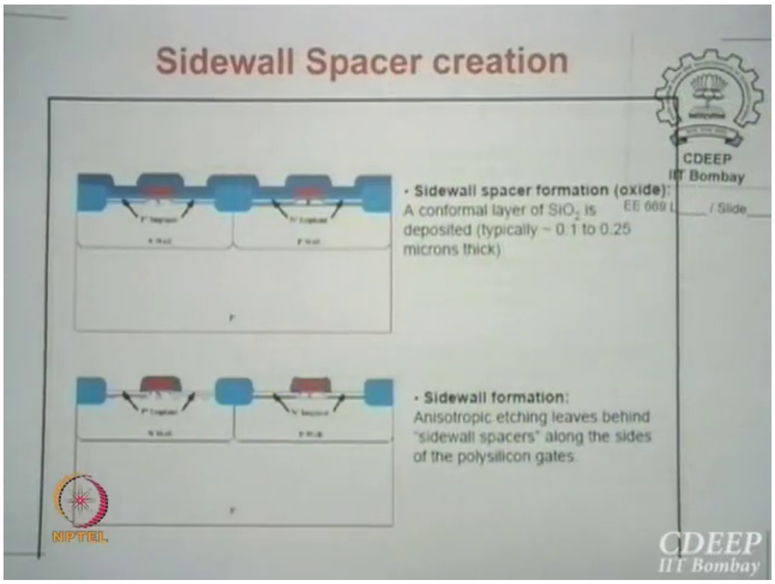
In 5 micron process or 3 microns which I worked in 70s, I never used to think this because the channel length was 5 micron, voltages were 5 volts and everything was well within my control. So there is another implant has been done, now one interesting feature you should see, poly open okay, the resistor is only on this side so whatever implant you are going to do now will also get implanted into poly, poly is open for implant. However, this is very lightly doped so poly is not really changing its property, it is still highly doped highly resistive because which is this implant, it should be lighter implant, why? I want the ability to be smaller in number because I want lightly doped so I actually make a well P regon here sorry N region there which is Light because N channel device you have in P-well so you made N-implant and what is important is this you must say, implant will always follow the edges okay implant will follow always ions it will go straight, so they will come here they will come here so this is called self-alignment.

What is the self alignment? Let us say during the last mask when this gate was delineated this shifted left but the next implant will follow the edges so source drain will be always the available to you only difference will happen is the region areas may change okay but at least transistor will always.... In our earlier process we used to 1st make source-drain and then use to make gate on that. So I must overlap my gate on source-drain so that otherwise if there is some area where there is no gate oxide so this is called self-alignment, wherever gate is there from the edges impurities will go down okay ions will go down that is the biggest advantage, this process is called self-alignment, there is this registration error this may shift all of them.

But so all the implants will still will follow the edges of whatever poly it has and therefore it will always gate will be always aligned with the source drain is that clear? What is the biggest advantage that provides the capacitance at the corners is smallest there and parasitic capacitance production is speed improvement okay so that is a very crucial requirement for high-speed circuits okay. Having done LDD for N-channel device what should I have do for now? Use the opposite mask or opposite resistor whatever it is and do same thing for P-channel device in the N-well so there is a thin P regions created which are going to be source drain, but these are lightly doped, what should the resistance of the resistivity of the actual source drain? It should be very low because I want a contact resistance to be low but this is very high so I cannot use this as a source drain is that point clear?

This is only lighter dope, the problem starts now this is where that word spacer came, I want only at the edges okay doping to be light but the rest regions I want it to be the heavier doped source-drain okay. So I must now protect some regions here where N-+ or P-+ diffusion implants will not come, have you got the point? This is an let us say P so if I protect some part sideways and then do implantation heavy implants then the small portion which is called spacer below that heavier implant cannot go and below is only lighter implants okay. So lighter drain or source can be created and the rest region I can create heavier dopes source and drain, this means spacer is required to create an LDD, this may be last slide for the day and then we will come back again.

(Refer Slide Time: 76:47)



So now to create a spacer and from where the spacer should be? From the sidewall because you know that is one which will block there next implant so I want extra outside covering gate which spacer covers actually gate region all around okay and then I have implant that area will not receive the next implant below that it will be only lightly doped regions okay so spacer is required for creation of LDD structures. What is the disadvantage of spacer? Any spacer and when you put LDD, the resistance of source drain will increase, is that clear? So spacer thickness is very very crucial how much R I would be tolerating at LDD regions is a very crucial effect because remember channel resistance will be the smallest, why? (())(77:45) you are pushing huge charge on the gate, you have(())(77:50) but still source-drain N+, P+ they are also very low but which is the maximum resistance will come, the below region of spacer.

So spacer width are very very crucial in decision of the speeds of the device is that clear? Typically if you want I may show you there will be ion N + let us say R n then R channel then R n and then R n +, this is drain this is source. This is very low, these 2 are also very low heavily doped okay but this LDD parts which is decided by spacer width okay they may decide your actual speed of the circuit okay. So the spacer has advantage of doing what? The removing short channel effect, but it has created compensation of speed okay so one has to keep varying how much speed that is why 10 gigahertz is not able, you know we try something I actually do not reach there okay. I had to give speed if I want to improve device, 1st I must have LDD why because if device does not work then why are we doing all this so device has to work so LDD I need okay.

So something I must too tricky which will reduce the spacer thickness I mean whatever doping there. Can I slightly change this resistance LDD? At the edge I do not want to change but slightly closer or below I can add something, how can I add I will show you a figure, this is my source drain and this upper portion is LDD, I want the below portion should be heavily doped that means (())(79:55) resistance to above okay. If I do a tilted implant I can go below so hollow has to be created okay, so that is how we can adjust our resistances but hollow is required on mask and also another damage at the spacer, spacer if it does not stop it will and it will also create its own problems okay. Spacer thickness cannot be uniform, its variation is 10 percent and that 10 percent variation actually changes 100% speed variation okay see you then next Saturday.