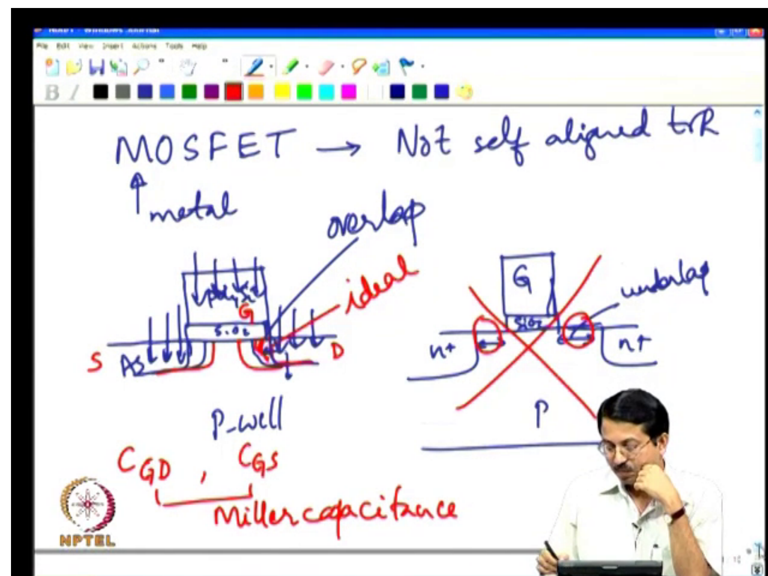
Nanoelectronics: Devices and Materials
Prof. Navakanta Bhat
Centre for Nano Science and Engineering
Indian Institute of Science, Bangalore

Lecture - 10
Metal Gate Transistor

Today we will look at metal gate transistors.

(Refer Slide Time: 00:24)



You may recall that the very abbreviation MOSFET M here really stands for metal because that is how we sort of building FETs, right. In the very beginning, the FETs that we used to build had metal gates, but the metal gate was due to certain constraints. We were not able to really do what we call in recent past polysilicon gate transistor and hence, we were building metal gate transistors.

So, when we were building this metal gate field effect transistor way back in 60s and you know maybe 70s, the problem with the metal was the following, right. The metal gate transistor was, it was not a self aligned transistor. So, what I mean by the self-alignment is that if you recall the MOSFET flow that we have discussed in one of our earlier lectures, if you recall I have mentioned that in polysilicon gate transistors that we have today have this gate stack which includes Si O 2 or of course, high k and polysilicon and here you do ion implantation, right. You convert you know polysilicon into n plus or p

plus that is how you make it a very highly doped polysilicon material and that is done along with the sourced in implant, correct.

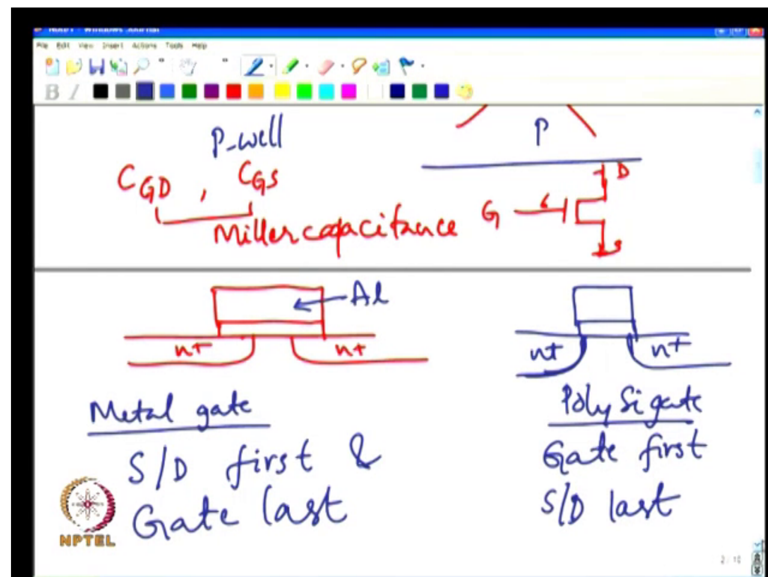
You see polysilicon was deposited as intrinsic polysilicon, but subsequently if it were to be n mos region, I implant arsenic and the arsenic will sort of come in here, right. This would have been p. Well, region for n channel transistor and arsenic would have also come in here and in this region; arsenic would have come in and stuck in the gate region, right. That is how I would have converted that gate into n plus doped region. Now, notice here that when you done with all these junctions, you see the junction is sort of self-aligned to the edge of the gate and of course, subsequently there is going to be some annealing process and because of the annealing, this dopens will diffuse in and also diffuse little bit inside laterally, but they are more or less aligned to the edge of the gate.

You do not want to have an unaligned or you know a transistor where gate does not overlaps source and drain. You know it is as long as gate is, right edge on with source and drain, but of course you do not want to have a transistor structure which looks something like this. You know that would be suicidal for transistor, right. Let say if this is your gate and this is your Si O 2 and only n plus region is here and n plus region is here, this is all p regions. When you are trying to do this, you somehow end up with this region which we sometimes call underlap as oppose to having something like this, where source and drain are actually going beyond the gate edge. This region is what we call overlap, ok.

If you have underlap where you know gate does not cover some regions of the transistor and as a result of that you will not be able to create an invergent channel. Here your transistor will not be conducting; it will not be a good transistor. So, this is certainly a no, correct. You should never have this. You should either have an edge on which is good because you know right now here you have junctions and immediately this region can be inverted by the gate because the gatefeild is influencing your channel region, right. This is ideal, but if you have a little bit of overlap, but not well, if you have too much of overlap is there, any issue well there is an issue and that is to do with a parasitic capacitance issue. This issue here is with respect to a parasitic resistance. You know this is the region where you will not be able to turn on the transistor, huge resistance for the transistor whereas, here what would happen is that if you were to look at the transistor, we call this overlap capacitance.

You see here is G and here is D and here is S. You have this overlap undesired capacitance called C GD and C GS, ok. Sometimes these are also called as miller capacitances that are essentially a capacitance between your gate and the drain, especially typically sources at ground potential let say for a transistor, but especially when you want to use a transistor as an amplifier, your gate to drain parasitic capacitance should not be too high, ok. That will in turn impact the bandwidth of the transistor.

(Refer Slide Time: 06:13)



In other words, you see a schematic representation of transistor in a circuit is like this, right G D and S. Let us say I am building an amplifier, this parasitic capacitance from input to the output of the transistor is not good. So, ideally zero overlap is best you can get, but may be very difficult to get. So, if you cannot get 0, at least you can get something you know non-zero positive, but certainly not negative.

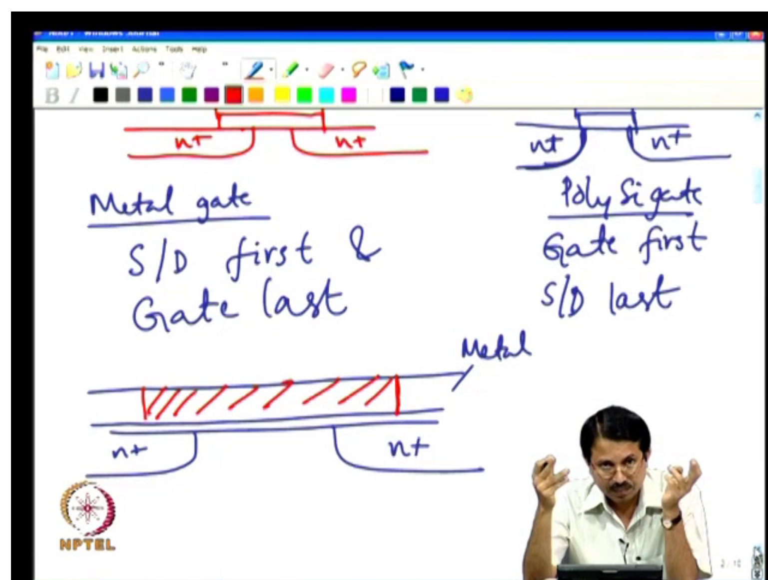
What happened with the metal gate? The very reason why we moved from metal gate transistors to polysilicone gate transistors was that the metal gate transistor invariably resulted in huge overlap here. So, if you are to compare your metal gate versus tran metal gate that we used to use in the past early days of the transistor technology, your metal gate transistor invariably looked something like this. You had this gate let say metal gate and a huge overlap, but still a working transistor. There is a parasitic capacitance, alright and on the other hand if it is polygate transistor, you get almost an ideal transistor structure just maybe a little bit of overlap here because of the self aligned nature.

So, what is not self aligned here you see if you are doing a metal gate transistor, first of all it turns out you cannot do source and drain implant after patterning your gate. You see let say that I am putting aluminum here. Let recall our transistor process flow in a polygate transistor. I first make the gates stack, right. This oxide and polysilicone gate is made and implantation of arsenic after implantation invariably. I will have to anneal at high temperature typically 1000 degree centigrade.

If you were to do the same thing here, what would happen at 1000 degrees? Aluminum will melt. Aluminum cannot withstand that high temperature even at 550. It will start melting. So, most of the metals that you know you want to use as gate materials, they would melt at temperatures which are used to anneal the source drain implantation and hence, in a metal gate transistor you do what is called source drain first and gate last process as oppose to here in a polysilicon gate transistor, you do gate first. Hence, you are defined gate stack you see and hence, I do not need anything else to define my source drain because it is going to be self-aligned, right source and drain last.

In this process, if you are doing source and drain first and gate last, you have already defined the junctions, correct and then, you grow the gate oxide, the process flow here will look something like this.

(Refer Slide Time: 09:30)



Junctions are defined, gate oxide is grown, gate electrode is deposited which is metal which is everywhere and now, all the gates are shorted, right. I need a photolithography

step to pattern the gate and that photolithography step should be aligned with already defined source drain junction. So, I am depending on photolithography step to do the alignment. So, what I do in order to avoid an underlap situation because when you are trying to align something which is already printed, there could be an offset. You see in order to avoid an offset, you build an extra margin. Even if there is an offset, you still get a working transistor which does not look like this, right.

So, as a result of that you do the patterning. When you do that patterning, you give an extra margin. You actually pattern a wider, a longer gate than that is required just to cover the sourced drain area, so that even if this gate shifts a little bit to the left you know, you still have an overlap here. Even if it shift to the right, you still has an overlap here. So, by giving an intentional margin I am going to get a working transistor because that is the first thing, right and then, you worry about all parasitics. So, the metal gate transistors always have the large overlap of gate over source and drain and hence, that is not a desirable thing.

This was one problem. The other problem with the metal gate transistor you know is also that you have metal to all transistors. It is the same metal for n channel, same metal for P channel let say, then the work function of the metal is fixed you see whereas, in case of a polysilicon gate transistor, your work function for NMOS transistor is closer to the conduction band because its n plus and your wow function for PMOS transistor is closer to the valence band.

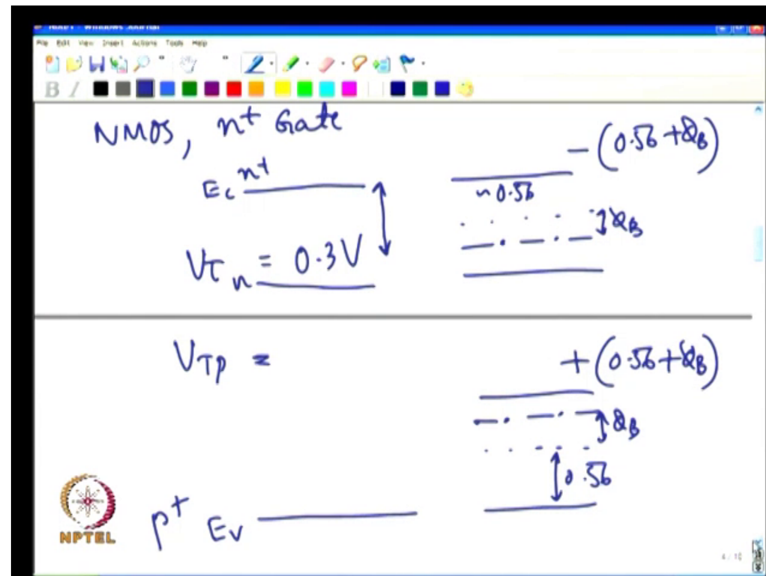
(Refer Slide Time: 11:48)

Now, why is it that important? It is very important because in order to get symmetric V TS for N and P, recall what VT equation is. Your VT is V FB plus 2 phi B plus Q D by COX let say, right

VFB ignoring all charges in the oxide, this is essentially phi MS which is work function difference, right which is phi MS plus 2 phi BQD by COX for polysilicon gate technology. What happens for NMOS, I have N plus gate? N plus gate meaning on the gate the Fermi level is very close to the conduction band edge. So, what is phi MS? As a result of that phi MS is negative in N channel transistor.

Remember that you have a vacuum level and this is the conduction band level on the gate. So, this 4.05 electron volt which is electron affinity and your substrate is P type, where Fermi level is below the mid-gap. So, work function on the P side is much larger. So, phi M minus phi S gives a negative quantity and in fact, phi M minus phi S is half the band gap plus phi B, where phi B is the bulk potential.

(Refer Slide Time: 13:27)



In the words, you see on the gate side, this is E c of n plus and on the substrate side, this is the mid-gap and this is phi B. This phi B is dependent on doping concentration, correct bulk potential. So, phi MS is really this which is half the band gap which is let say 0.56 and phi B. So, it is minus of 0.56 plus phi B that comes here. 2 phi B is a positive quantity, but little less than this because you see phi B can never be 0.56, right. Phi B is always less than 0.56. So, this minus this will be a small negative quantity and over powering that you will have another positive quantity and eventually end up with 0.2 0.3 whatever threshold voltage depending on your doping concentration and oxide thickness, right. This is for N channel.

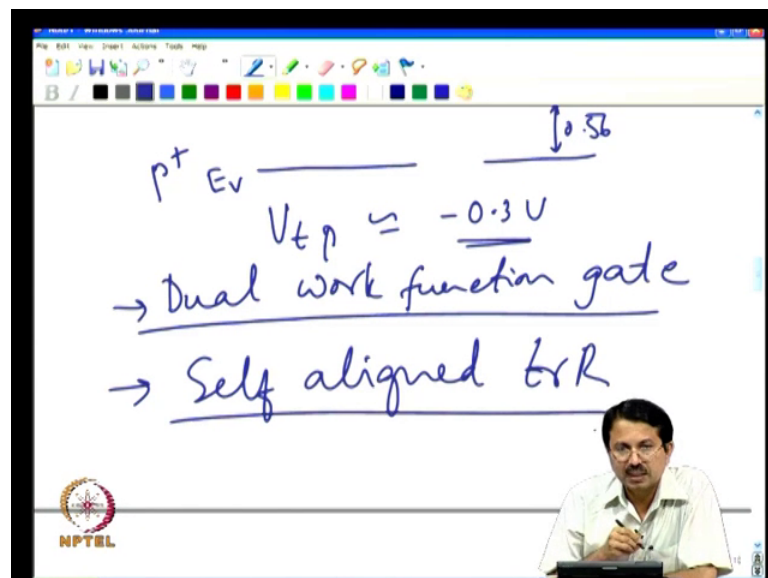
Now, P channel very interestingly let say if you do the right doping and all let say your VT of N is 0.3 volt. Now, you see what happens to P channel VTP is again the same equation except phi MS is going to be positive because now phi M for P channel transistor is P plus which is 4.05 plus full band gap because E V is here, correct. The Fermi level will be here for PMOS. What is the work function here? It is the phi B. Let me draw this, so that there is no confusion, right.

So, here this is E V valence band, right this is P plus Fermi level is on the gate close to the valence band, right. That is why we call it as P plus. On the other hand, substrate is N type, this is mid gap and this is your Fermi level and this is phi B and this is half the band gap 0.56. So, phi MS, phi M is larger than phi S. So, phi MS is a positive quantity,

correct and phi MS is positive 0.56 plus phi B. Now, what happens here the band bending is minor. I mean this is a negative quantity. Now, it is P channel transistor bands are bending opposite.

So, minus time 2 phi B is plus this still giving you a little small positive quantity and again you have this QD which is an opposite charge and that will also be a negative quantity and all that put together will give a negative VT which is almost same as this positive VT for N channel. That is excellent, right.

(Refer Slide Time: 16:14)

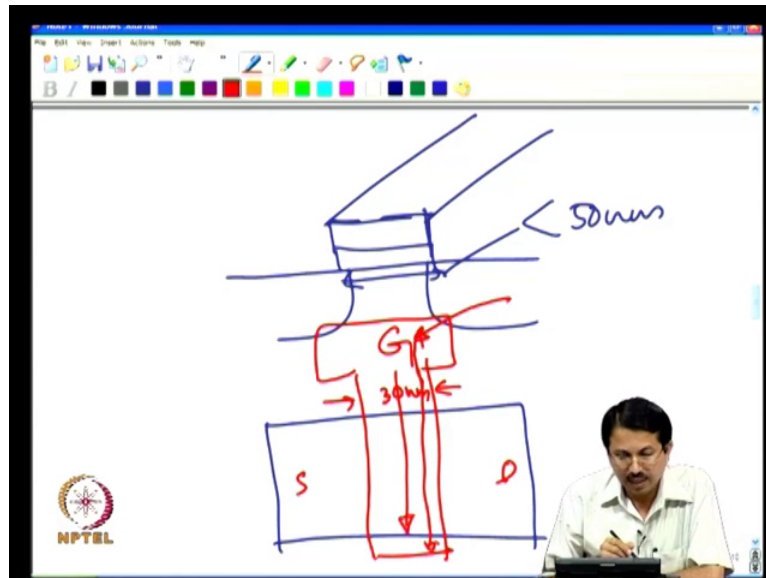


So, in other words, your V t for p will become equal to minus 0.3 volt. This happened simply because we had dual work function gate. This would never happened if you have identical gate for n channel and p channel. Let suppose that the gates for n channel were you know n plus for p channel, they were also n plus, then you will get a huge negative number for V t, right because this you will not have a positive quantity. This will also be a negative quantity because this is sitting here at n plus.

And similarly, you know you will not be able to get an appropriate V t matching for n and p channel transistor. So, this is the problem, right. And hence, the dual gear technology was excellent, right and as soon as we figured out how to you know do this polysilicon and dope polysilicon heavily n plus p plus, we migrated to the dual gate polysilicon technology, because you get this as well as self-aligned transistors, correct two very important metrics.

So, why are we talking about metal gate today? If this is why in the first place we moved from you know metal gate technology to polysilicon gate technology, right well there is a reason for that. The reason also becomes very clear to you if you look at what we have been doing so far, right.

(Refer Slide Time: 17:49)

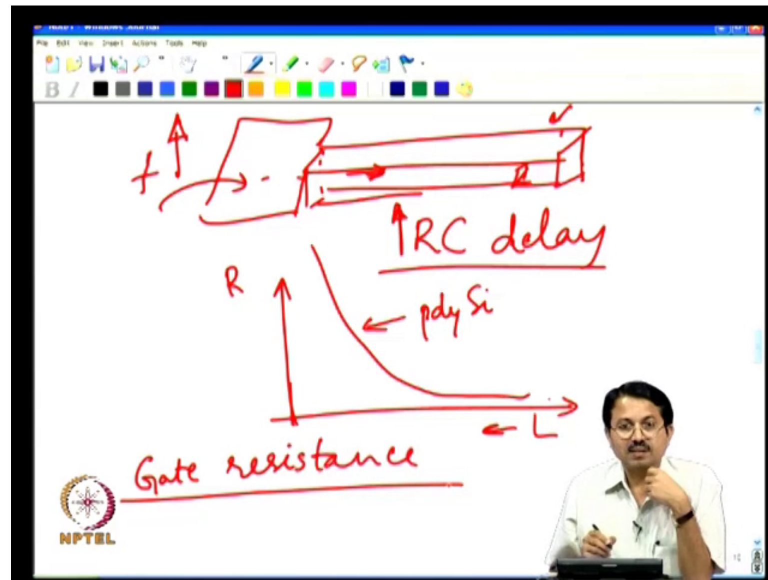


If you look at the transistor, you see this gate length this is my gate, right and this gate length has come down from in microns, 10 of microns. Today we are talking of less than 50 nanometer. Let us for a minute look at this direction, not the gate length direction, but the width direction. In other words, if you were to look at the top view of this transistor and let say this is the width and this is where you will have a gate contact, this will be gate and you will have source here and drain here and so on and so forth.

This is shrinking, right. This is shrinking to 30 nanometer and so on and so forth from where it was micron, but you see the widths in any circuit are dependent on how much current you want from the transistor, right. If you want more current, you scale the width and you increase the width, right. So, width is not necessarily in your control. You make the lengths as small as possible, so that transistor becomes fast transistor because the distance between source and drain is very small.

Now, if you for a minute look at in this direction, what will happen is the signal from a previous stage will come on to this contact and this signal should propagate along this direction correct.

(Refer Slide Time: 19:34)



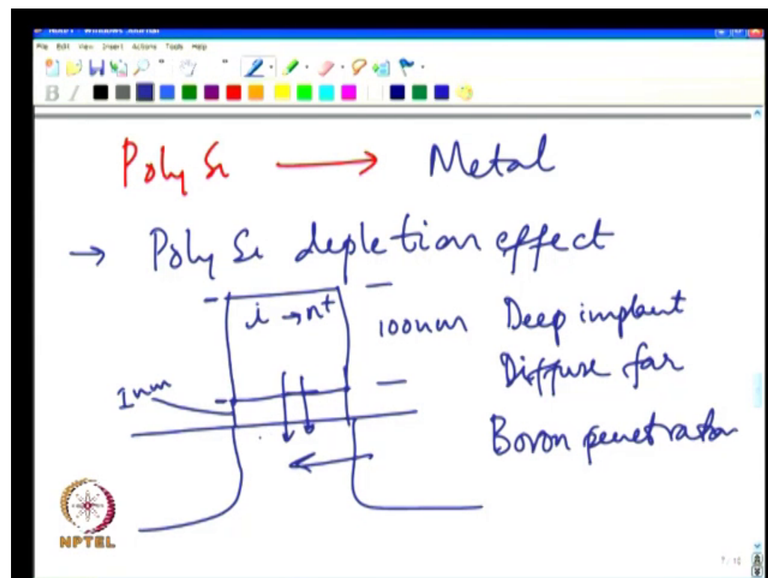
Now, let us look at that and how does it look you know that essentially looks like you know something like this, right. So, the signal is coming at one edge. Let say this is where the signal has come and the signal has to propagate along this and you see this is not a super conductor, right. I mean this is after all you know some material polysilicon. You have doped it very heavily, but none the less you still have some resistivity. Now, its resistivity could be a few ohm per micrometer, ohm per square kind of resistivity. So, what it means is that as we start shrinking this, what is that? That is channel length, right. As we are starting to shrink it that is one aspect what is happening is, simultaneously the frequencies are increasing, right.

We are no longer talking of mega hertz; we are talking of gigahertz, right. So, at this large frequency this will start acting as a transmission line because it has a finite resistance. The resistance is not you know 0 as I said and there is a capacitance obviously because that is a gate insulator capacitance. So, this is essentially RC delay, ok and this RC delay is not negligible for two reasons. As I mention again this R is increasing. Why is R increasing? If you were to look at your resistant, you have decreased the channel length, correct. Earlier when the channel length was long, there was; what is area of cross-section for the current flow here that is the thickness, correct and the length that is the area that is available, this length is decreasing.

So, area of cross-section is decreasing because of that and not only that, we have also scaled the gate stack. We have not been using very thick gates. We have also scaled that. So, even thickness has decreased, right. So, as a result of that this resistance has really gone up like this polysilicon gate, so because of this large resistance, the RC times have become comparable to reciprocal of the frequency of the signal of interest. I am using this microprocessor at you know few tones of gigahertz reciprocal of that is you know it depends on that will dictate how much RC delay you can tolerate, that is less than nano seconds, ok.

This is a serious problem, right. So, the problem is the gate resistance. Gate resistance has become a serious problem, right. When we say gate resistance, gate resistance as I mentioned for the signal to propagate you know from one end of the gate to the other end of the gate. What it means is that when the signal comes here, the transistor will slowly start turning on from this edge. So, current front moves from this edge and eventually when the signal reaches this, you will have whole transistor conducting which is not good, right.

(Refer Slide Time: 23:18)



Your current is much lower than what should have been for certain duration of RC delay of this gate, right and hence, we want to replace this polysilicon gate. What is the solution? You convert this material into something which has much lower resistivity.

What is that lower resistivity metal? This is one thing, but there was another equally important problem that came about that was to do with the following, right.

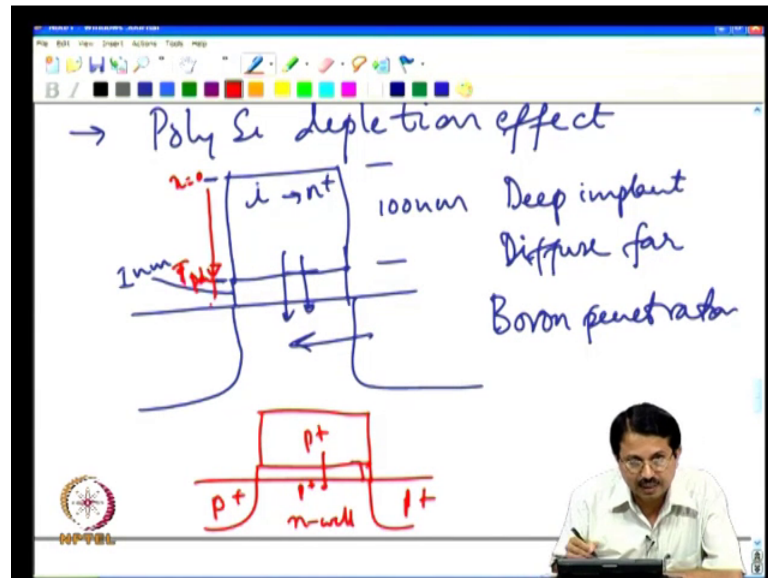
Remember polysilicon when it is deposited, it is intrinsic and it had to be converted into either n plus or p plus by doping. If you do not dope it, it will not have low resistivity. So, that resulted in one more problem which I will again briefly discuss and that is called polysilicon depletion effect, right. So, what it essentially means is that you have this polysilicon let say its 100 nanometer thick. You need to implant and diffuse, correct. Look at the fine balance that we have to strike here. How do you dope the polysilicon completely? You see polysilicon which is intrinsic to begin with I will call it I for intrinsic should be converted into n plus. For this to convert to n plus, this whole region from this point all the way to this point should have very heavy arsenic doping, ok.

How do you do that? One of the two ways, right either you implant very deep. If you cannot deep implant very deep, you diffuse. You have implanted only this point. Let do a huge doze implant, diffuse it down, so that it will come down, but the problem is the following, right. It is the same implant that is going to the gate and the source and drain. So, if you want to dope the whole thing, your source and junction could be at least as deep as this because it is exactly same process. It is not going to be a different process that is the beauty, right. It is a self-aligned transistor. It is the same process that is defining it n plus and n plus source drain, but deep source drains are not good because they would in turn result in bad short channel effect, they will and so and so forth.

We have talked about this coupling volume has to decrease, correct. That is what we have discussed quite extensively. That is one thing. The other thing that happened was sometimes when you actually did this somehow, let say you did this. This oxide that you had was ultra thin, right. One nanometer is what we were talking about Si O 2. When you have such a thin oxide, you know these dopens can also diffuse through the oxide especially. So, with boron because boron is so light as oppose to arsenic. In fact, in literature you will see this problem called boron penetration. What it means is that the boron can easily penetrate through this ultrathin 1 nanometer. It is as if it is non-existent there, ok.

So, now you do not want the boron to come. Boron will be in p channel transistor. In n channel transistor we can talk of arsenic penetration, but that is not bigger problem, but in p channel transistor, it is compliment of this as you know, right.

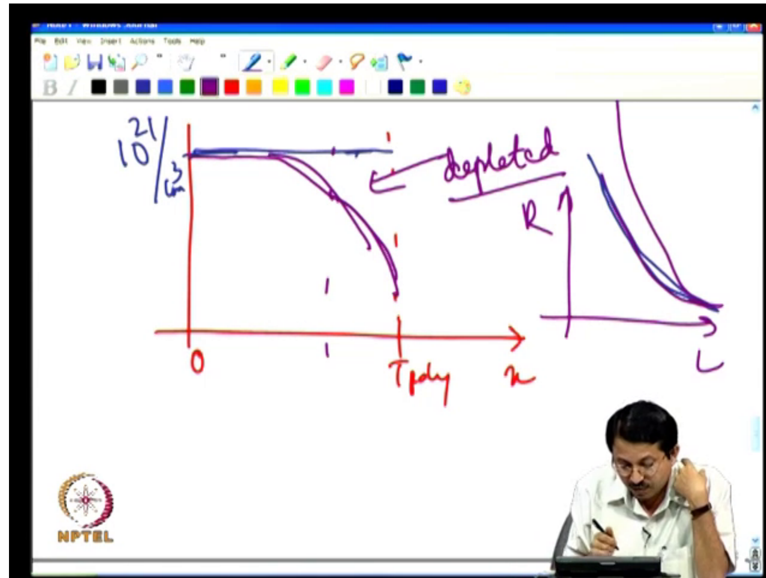
(Refer Slide Time: 27:07)



So, then we are talking of this is made p plus. This is made p plus and your gate is also made p plus. This boron will come here that will make this p plus that is not good. You do not want boron to come in the channel, correct. You have actually a n. Well, here you do not want that n to convert into p type you see.

So, invariably you had to end up with a situation where the polysilicon very close to gate oxide will not be as heavily doped. In other words, if I were to look at the doping profile, let me call it you know this is what I call x axis, right where this is what I have been calling y axis. Let me call it you know x equal to 0 here for just the sake of this plot and this is t poly which indicates polysilicon thickness. So, then what you probably will end up is something like this.

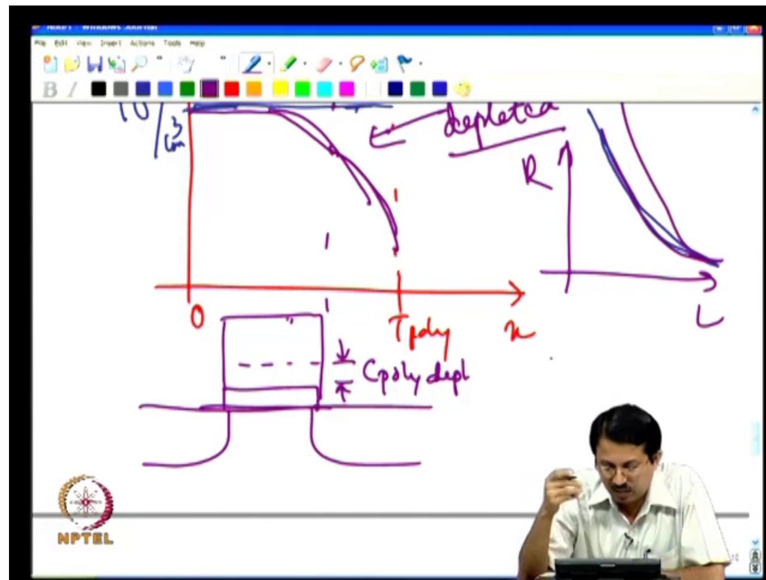
(Refer Slide Time: 28:04)



X equal to 0, here x equal to t poly ideally, correct. This is where your gate oxide starts. Remember that ideally you want this whole thing to be doped very heavy let say 10 to 21 per centimeter cube arsenic or boron, but I cannot do that, right for one thing. The junctions could be deep. F,or the other thing you can end up with having this penetration. So, invariably what you will probably end up is something like this because you are afraid to push it too far.

So, the region very close to your gate oxide will be depleted and this is what we mean by poly silicone. Depletion is not as many dopens. What is the impact of this? The impact is two-fold. One, obviously the gate resistance will increase even more because the plot that I had shown you earlier of R versus l is assuming that whole polysilicon is doped, right. Let say that corresponds to this case, but now if you have this case, you may have something like this because this region is not doped you see and hence, resistivity will increase in that region that is one, but there is even more serious problem with this and that is if you look at the vertical direction, ok.

(Refer Slide Time: 29:39)

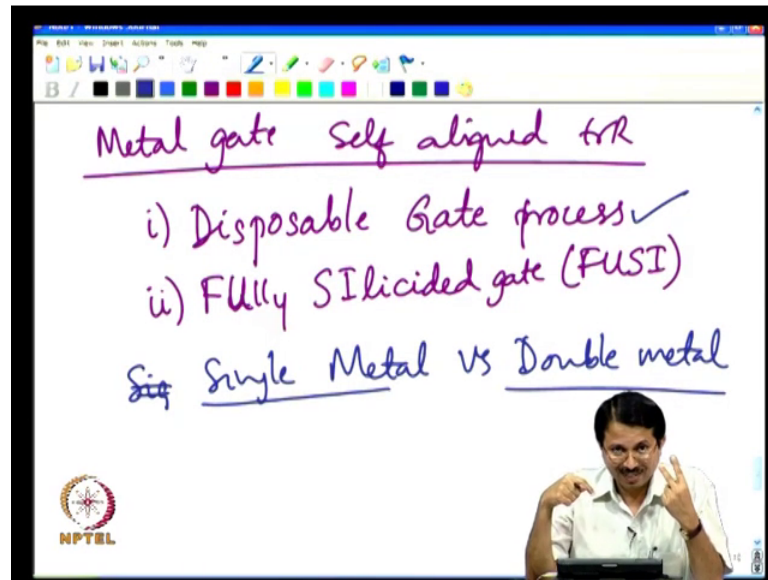


What happens now in the vertical direction is this is your gate, you see this polysilicon. We used to only consider the gate oxide capacitance here and the silicon capacitance assuming that the whole thing is like a conductor, but now there is huge depletion here which means there will be effectively another series capacitance.

So, this is what is called polysilicon depletion capacitance. So, whenever a new capacitance comes in series, the total capacitance goes down as oppose to resistancy series. So, what it means? It means is that your gate coupling to the channel degrades. Your gate cannot effectively couple to the channel because now it is C o x and cc depletion that will certainly be much lower than C o x. So, your transistor will have lower current, right. So, these were real problems you know. We were stuck with these problems and that is why we had to change to a metal gate technology.

So, just to recap that you know the resistance which increase RC delay was not desirable and poly depletion effect was another undesirable effect, right. Both this forced us to change polysilicon to metal gate, but now the challenge for now is how do you make a metal gate transistor which is self-aligned because we have been doing self-aligned transistor. We do not want to go back to our old transistor which has a huge overlap because now we are talking of huge frequencies of operation. So, I must get self-aligned transistors that are the challenge, right.

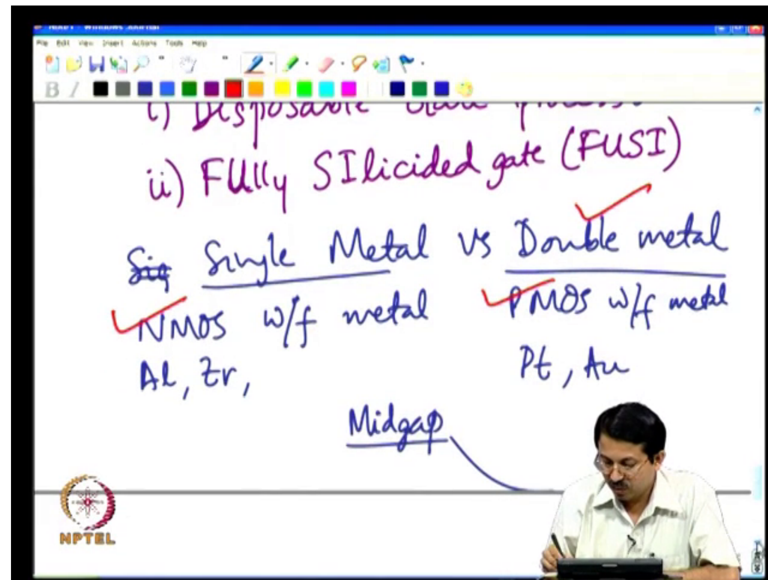
(Refer Slide Time: 31:35)



So, I want to get metal gate self-aligned transistor. So, there are you know possibly two ways you know. One can get there is what is called a disposable gate process and there is another which is called abbreviated as FUSI fully silicided gate, ok

Of course, this is more complicated process and today we have in manufacturing all 65. 45 nanometer technology are actually using disposable gate process. We will see what disposable gate process is in a minute. Then, the first question to ask is do you want to do a single metal gate or double metal gate. Now, this is important consideration. Remember I told you dual polysilicon gate technology. Gate was excellent because you got n plus kind of a work function on NMOS on the gate side and p plus kind of a work function for PMOS on the gate side, correct which was dual work function. So, if I need to have matching V TS, I must have two metal gates, right. So, we call them you know NMOS you know NMOS work function metals and PMOS work function metals.

(Refer Slide Time: 33:35)

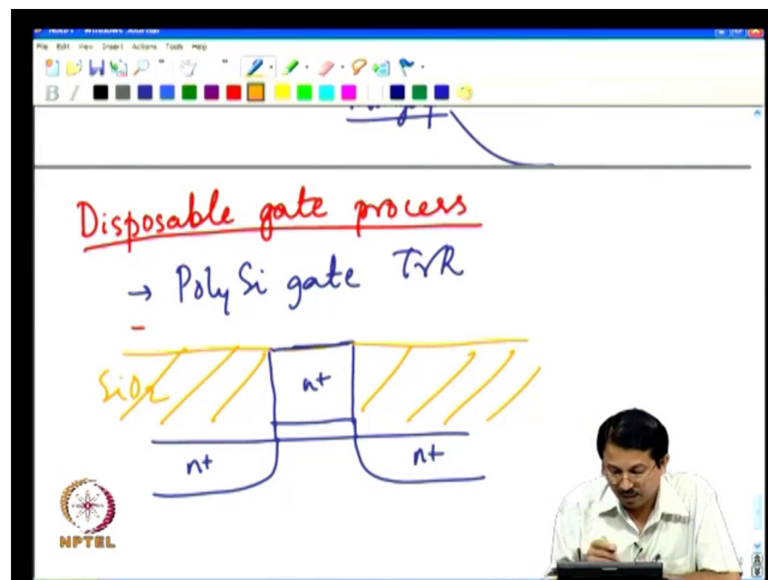


So, you see metals have different work functions depending on which metals you are looking at. For example, if you look at aluminum, zirconium, they all have low work function. Aluminum is 4.1, zirconium is about 4.2. That is very close to n plus because n plus is 4.05 that is the electron affinity for silicon conduction band, right n plus that is where the work Fermi level is. There are others such as platinum and you know few others you know gold also has large work function and few other work metals are there like they are all 5.2, 5.3, 5.1 which is what is there in p plus because p plus is 4.05 electron affinity plus one band gap which is 1.1, correct. So, the p plus work function should be like 5.2 E v. So, there are different kinds of metals that are available.

There was also a proposal for a while is that look if I had to use two metals as oppose to polysilicon. Polysilicon was an intrinsic polysicon which was deposited and then, converted to n plus during n channel implant and p plus during p channel implant, but here I need to have two different deposition process and two different lithography process you know. I need to first deposit aluminum and etch it off in PMOs regions and retain only in NMOs regions and then, block NMOS region. Deposit some other metal and you know again pattern it more complicated, right which require two deposition, two etching, two lithography to pattern the gate as oppose to a single deposition and single etching process and there was a proposal called you know mid-gap metal gate. There are some metals you know which have work function neither 4.1 nor like 5.2, somewhere in between like you know 4.6, 4.7 E v.

So, the idea was that you know you have a single metal gate whose work function is at the middle of the band gap, right. So, you compromise both n channel and p channel transistors identically. So, you would probably still be able to get matching V t, but V t will certainly be compromised. So, that was not really usable solution. So, what we have been doing today is double metal gate using disposable gate process technique and double metal gate one with NMOS and one with PMOS work function. As you know the work function is very important because V t will not be obtained properly otherwise, ok

(Refer Slide Time: 36:56)



So, now how do we do this now? So, let us look at disposable gate process. So, what we do here is that we just go through the usual process of making a polysilicon gate transistor that is you have n well or p well. Whatever it is, you actually do a gate oxide, do polysilicon deposition although I want to get metal gate right, but for the time being I do polysilicon and pattern polysilicon wherever I want to have the gates on n channel region as well as p channel regions, then do the source drain implant. When I do the source drain implant, polysilicon is sitting here, correct. So, it will still give me self-aligned transistor, correct. So, in other words, I will get let say n plus n plus and this also becomes n plus. Let say now what I will do is the similar thing will be done for p plus also, ok.

Now, after this what I do is that I deposit an insulator on top everywhere. Let say you deposit extra insulator and then, etch it off. Let say I deposit Si O 2 and polish this Si O

2, such that I will end up with all Si O 2 in this region and this region and in this region. Now, this gate is exposed. So, what will happen here then is that I guess, right. So, this is Si O 2 here, correct. Similarly, in p channel regions also. Now, what I do is I have Si O 2 here and I have silicon here.

(Refer Slide Time: 39:47)

I can use h chemistry and I can get rid of this polysilicon gate, remove it completely. So, when you remove it, you will end up with a cavity there, correct. Remember sourced and drains are already formed and they are formed self-aligned to this edge, correct and the from self-aligned to this edge because this polysilicon was a place holder. It was just sitting there to define my self-aligned transistor. Now, the polysilicons job is done because I do not want to use it eventually in my circuit because it will have a huge resistance. So, etch it off, etch it off selectively by creating a structure like this.

Now, fill this region with whatever metal that you want in n channel region let say aluminum, right. Of course, you know you do a lithography. So, make sure that it does not go to p region. All your n regions will get this and similarly, you do another set of process meaning only the deposition opening and cavity formation is already done. It is done at the same time anyway, right. You know you have another cavity and for p channel transistor, this is p channel transistor p plus p plus and there is a cavity here. This is Si O 2 and this is also Si O 2, this is also Si O 2 and you fill this cavity with a different

metal. This will be PMOS work function metal and this will be NMOS work function metal.

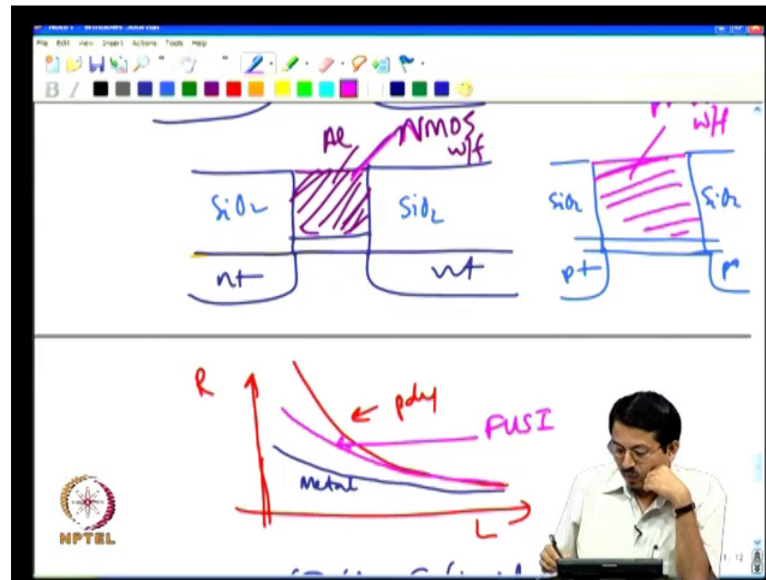
It is more complex process that is true because you know you have to etch, had the polysilicon, had to etch it away, fill it differently in n region, fill it well, you deposit everywhere, but you know retain it here pattern lithography and retain it here and similarly here and and so on and so forth, right.

So, I will just take back what I said you do not do lithography here. You just do chemical mechanical polishing, right. I mean you do not necessarily do lithography, but none the less you know you can do this process such that metal is there and source drain was before metal. So, you do not have to anneal it at high temperature, correct. So, there is no danger of metal melting, but because I add that dummy polysilicone, I got self-aligned transistor and hence, the name disposable gate process. There was a polysilicon gate and that was only a place holder and that was disposed off subsequently. That is how I got this metal gate here, metal gate here perfectly self-aligned transistor, right. This is one way of doing that and this is done very extensively today.

Well, there are lots of issues you know sort of filled in that region so easily with you know the marker pen here, but really you know you will have to make sure that there is a huge cavity here. The metal should deposit conformally everywhere and you do not want to end up with the cavity here where metal is not deposited properly. So, there is a lot process optimization that goes in, but you have to do that because there is no other way you can build transistors today.

If you are talking of transistors less than 45 nanometer, you take for granted that the process is certainly more complex, but that is worth it only. When you do all that complicated process, you will be able to you know get a very well behaved transistor. So, what is the implication of this you know we started off with the main problem that we had, right.

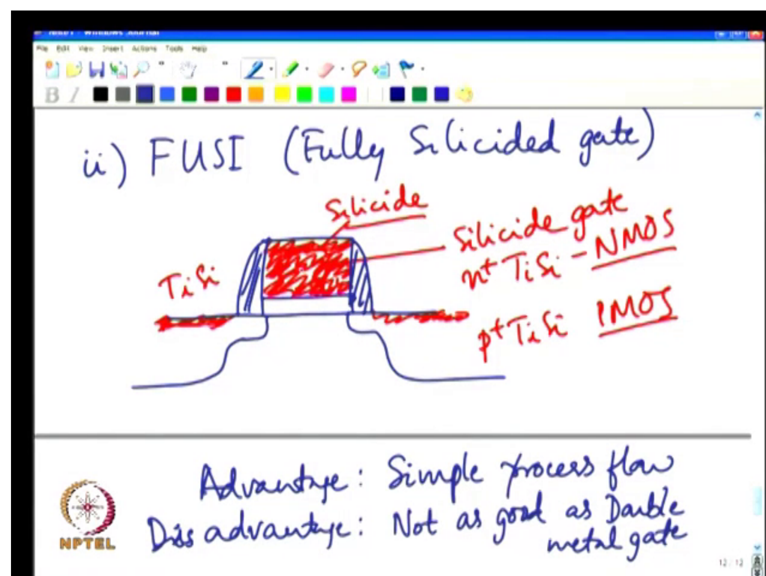
(Refer Slide Time: 43:55)



That is my R versus length you know. It looked like this you see if I go to metal gate, it would look like this. So, this is the metal gate whereas, this is the polysilicon gate especially where your shrinking length very significantly, metal gate will outperform polysilicon gate very significantly.

What is the second option? Second option is called FUSI or fully silicided gate, fine. So, what do we do here as the name suggests, the idea is to exploit silicidation process to create somewhat like a metal gate transistor.

(Refer Slide Time: 45:19)



What is a silicidation process? If you remember I had explained this to you earlier in one of my lectures and that is I told you that your transistor looks like this. There will be a spacer here, there will be a spacer here and eventually the source and ground will come in contact here. You will have a shallow extension here and then, a deep sourced drain here, correct. I told you that this region we convert them into a combination of metal and silicon or silicide. You deposit a metal, anneal it at high temperature and there is a reaction between metal and silicon and you end up with a silicide, correct.

So, here this region and this region these all gates converted to silicide. So, the idea of fully silicided technology is that you let say I am doing a titanium silicide, you convert this whole region into titanium silicide. So, strictly speaking we should not call it metal gate. It is more like a silicide gate, right. In other words, it is a combination of titanium and silicon which is either n plus doped or p plus doped. Very interestingly when you convert a silicide, the work function of this silicide tisilicide is governed by what is your initial work function of silicon, right whether you had an n plus silicon or p plus silicon.

So, in all n channel regions, you would have n plus gates, correct and that becomes an n plus tisilicide and in p channel region, you have p plus tisilicide because you had a boron heavy doping, right. Here you had a arsenic heavy doping. So, as a result of that you have been able to get somewhat like a metal gate. It is not quite metal gate because this is not comparable to having an ideal metal gate, right. In other words, here if you were to interpose another curve, what you will probably see is that you know you may get something like this for a FUSI fully silicided technology.

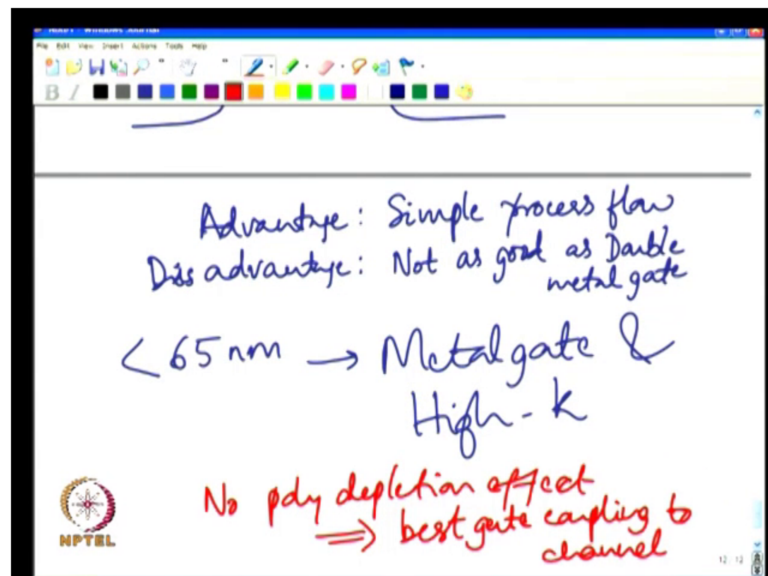
Fully sided silicided technology is certainly better than polysilicon gate. Certainly it is not as good as a metal gate. What is the advantage here? The advantage is simply that the process is not very complex. You see I do not have to do a disposing of the gate, etching of the gate and then, doing another deposition for n like gate and another deposition for p like gate and so and so forth, right.

So, the advantage of a FUSI process, right simple process, but there is a disadvantage also, right. As you know double metal gate that we are talking about and also the whole thing has to be silicided you know. If you cannot silicide, if part of that silicon is left with, then that is also not a good thing, right because that region will will be high

resistivity region, right. So, your total resistance will also increase because of that region, right.

So, as a result of that you know fully silicided gate is a theoretical possibility. You will see that lot of literature exist where in people have demonstrated it as doable and you get reasonable threshold voltage matching for n and p because it works as a dual metal function work function gate because it does not improve RC delay as much as a double metal gate technology. You know it is not really a practically implementable technology. So, that is what we essentially do with respect to metal gate technology.

(Refer Slide Time: 50:27)



In fact, today all technology, today the state of our technology like as I said you know anything less than 65 nanometers, CMOS technology is a metal gate and high k as we discussed in the last lectures, where we also have a high k gate dielectric, right. High k gate dielectric and metal gates sort of go hand in hand. You know they happen simultaneously in terms of replacing the silicon oxide as a gate insulating material and also in terms of replacing polysilicon gate as your gate electrode, ok.

The best thing about a metal gate is that there is absolutely no polysilicon deletion issue apart from the resistance issue that we talked about earlier RC delay and all that, right. So, the polydepletion effect is completely eliminated, right. You get the best gate coupling you know which means the best gate coupling to channel and hence, invariably you know you will see that your high k with metal gate as oppose to polysilicon gate will

certainly have much better on current of a transistor as oppose to polysilicon gate transistor, fine.

So, well let me then summarize the discussions that we had today. You know MOSFET as the name suggests is you know a metal gate FET that is how we started building FET, but when we started building an FET with metal gate, we had to resort to a gate last process and sourced drain first process because otherwise metal would melt during the source drain annealing process, right and hence, the gate had to be defined using photolithography process and you ended up with huge overlap, gate overlapping source and drain and parasitic capacitances and that is how we sort of quickly migrated to polysilicon gate technology. Although we continue to call them as MOSFET, M here was a misnomer really for last 3-4 decades. M was misnomer because it was not really metal gate; it was a polysilicon gate technology highly doped polysilicon.

The greatest advantage then was most importantly you had self-aligned transistor and you also had this you know n plus and p plus kind of a dual work functions for your n channel and p channel transistor which resulted in very well mashed threshold voltages for NMOS and PMOS, but as we started shrinking the gate length you know in a matter of 65 nanometer, the resistivity of the gate started really influencing RC delay for the propagation through the gate electrode and that is when we were forced to replace the polysilicon gate with the metal gate, but then the challenge was really to get the metal gate using a self-aligned transistor flow, right.

That is how we sort of resorted to having a disposable gate process wherein we continued the best practices of polysilicon gate technology wherein polysilicon was just a place holder to define that self-aligned transistor subsequently owed the polysilicon and you defined different metal gates for n channel and p channel regions. Hence, you get the self-aligned transistor as well as the very well mashed V TS because I am using the double metal gate process, ok.

So, we will stop here and we will continue in the next class.