Design and Analysis of VLSI Subsystems
Dr. Madhav Rao
Department of Electronics and Communication Engineering
Indian Institute of Technology, Bangalore

Lecture - 57
Introduction to Repeaters in Interconnect Engineering

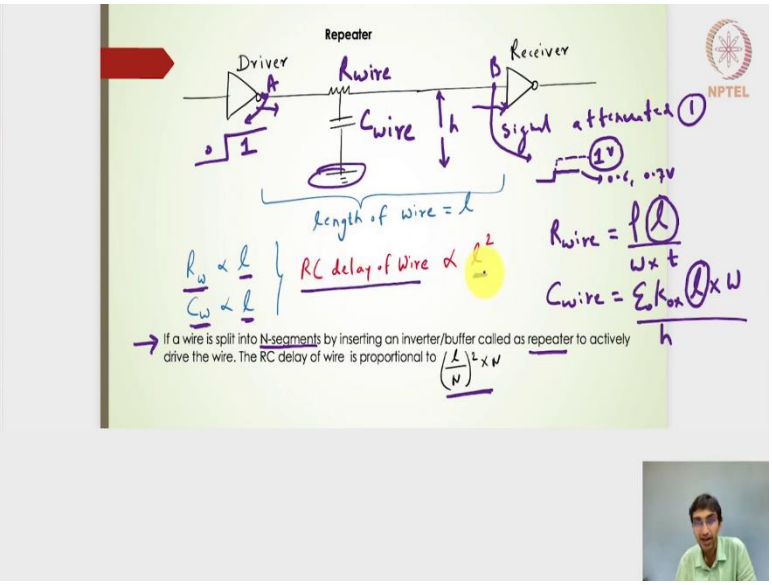
Hello students, welcome to this lecture on Interconnects. In this particular part of the lecture, we will start with the repeaters, importance of the repeaters in a long interconnect wire and then proceed further to see how do we design the repeaters and how many number of repeaters we will require, so as to optimize what is the optimized number of repeaters that is required to minimize the overall delay.

Later on we will see what happens to the characteristic energy properties for a long wire with the repeaters and without the repeaters. Repeaters in general are nothing but an inverter or a pair of inverters, so that your overall signal strength while it reaches to the other end of the wire, it is still the same signal strength as that of where the signal has started or propagated from.

You can consider, if I have one inverter and then another inverter and it is kind of connected by a very long wire, then the output of the first inverter, when there is a signal at the output of the first inverter and if it has to propagate along this long interconnect or a long wire, till the time it reaches to another part of the inverter, the signal strength might actually get attenuated.

So, we add some inverter, so that the signal strength improves and we have relatively a stronger signal, while it is reaching at the other end of the wire, that is just a brief introduction.

(Refer Slide Time: 02:03)



Let us you know start looking into more details, this is what I had said earlier we have a driver circuit which is in this particular case it is an inverter. Then we have a receiver circuit, again it is drawn an inverter here and then let us say that we have a long wire here. If the signal starts here, the output of the driver inverter circuit and by the time it reaches here by a long wire, the signal might actually get attenuated. In that sense if I draw that signal, you know you have 0 to 1 here.

At the output of the driver circuit I have 0 to 1. At the input of this receiver signal, if I have a really a long wire, then I think the 1 voltage may not be 1 voltage. This is actually the one value, but here I will get an attenuated signal, it could be 0.6 volts or it could be 0.7 volts and so on. There will be an attenuation based on the length of the wire, the there will be a signal loss.

In that sense we need to add some kind of a repeater circuit, so that you get the expected signal level. Now, here in this particular circuit in this particular schematic what I have shown is a long wire connecting the two inverters, one is the driver and another one is the receiver. I have stated that this is the length of the wire, so l is the length of the wire.

The wire or an interconnect can always be represented in the form of an RC circuit. A distributed R and then a distributed capacitance is here and then this is of the wire. I am going to write it as R_w and then C_w. Now, this w and then C_w is nothing but rewrite this

as R_{wire} and then C_{wire}. C_{wire} is nothing but the overall capacitance for the long length of the wire.

If I consider the R_{wire} by definition it is nothing but the resistivity of the wire, in this case it will be,

$$R_{wire} = \frac{\rho l}{wxt}$$

Here it is actually $R_w \propto l$. C_{wire} if I go by definition it is nothing but,

$$C_{wire} = \frac{\varepsilon_o k_{ox} lw}{h}$$

Where, h is representing the spacing between the two wires.

If I consider this to be a ground plane, then in that sense I will have whatever is the spacing will give me. Spacing between this particular long wire and then this particular ground plane whatever is the spacing here will give me the capacitance.

Here also the capacitance parameter is also defined by the length of the wire. If I am going to evaluate the overall delay, the overall RC delay of the wire, $R_w \propto l$, $C_w \propto l$. RC delay of wire $\propto l^2$, what it really means? It means that as I increase my length of the wire, the overall delay from this particular point A to this particular point B.

The overall delay for the wire for the signal to propagate from the A part of the wire to the B part of the wire, A is directly proportional to the length of the wire. If I increase my length of the wire, the overall delay is likely naturally to increase by the square of the length of the wire. If the length of the wire is 1 mm, I will have a directly proportional to the 1 mm square, if tomorrow the length of the wire is actually 10 mm, it is going to be 100 times more than that what we had acquired for the length of 1 mm.

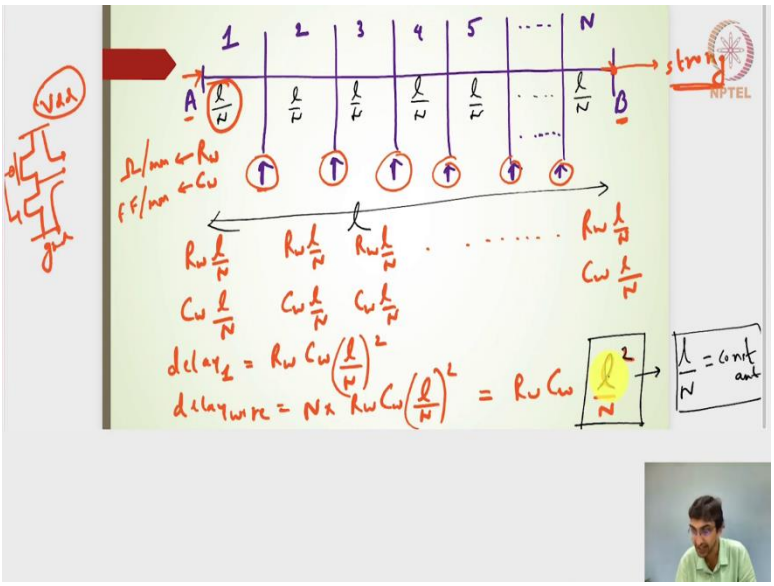
That is something we have to look into it, because we do not want the interconnect propagation delay or the delay due to the interconnect or the delay due to the wiring to be very very large then that of the computational delay.

The driver and receiver it could be substituted some other combinational circuits than that of whatever I have shown here, instead of the inverter, it could be any kind of a

combinational circuit. If that is the case, the computational delay whatever we will achieve and that should not be dominated by this overall length of the wire.

We need to do something about it. One of the statements I have written in the slide here is, if a wire is split into N segments by inserting an inverter or a buffer called as repeater to actively drive the wire. The RC delay of the wire is proportional to the $\left(\frac{1}{N}\right)^2$ x N. Let us have a look at it. There are two problems here, one problem is the attenuation problem. This is one problem if I have a long interconnect or a long wire. The second problem is the RC delay of the wire is directly proportional to the square of the length of the wire.

(Refer Slide Time: 07:37)



Let us once again look into this long wire. Let us say that I have a long wire and then this is A and then this is B. At A I will have an a driver circuit and at B I will have a receiver circuit. Let us say I am going to add a repeater, so that it is attenuation gets improved.

While it reaches to the point B, it will have the same signal strength as that of the point A. I am going to add repeater and I am going to add N - 1 repeater, so that I will actually segment the wire into N segments. I am going to add a repeater here, I am going to add a repeater here, repeater here, repeater here, here and here. I am going to say that this is dot dot dot, I will have this.

I will make to divide the wire into the segments. This is segment number 1, this is segment number 2, this is segment number 4, 5 dot dot dot and then this is my Nth segment. I have

divided the long wire into N segments. Let us say if it is equally divided, then I will have the length as l/N, the length as l/N and the length as l/N, length as l/N and I will have dot dot dot and then length by l/N.

Let us say that this overall length of the wire is l. If I separate the long length of the wire into N segments equally then I will get each segments length as l/N. Let us say that the long wire is also characterized by R_w and then C_w per unit length of the wire. What it means is, R_w is a resistance per unit length of the wire, the units will be Ω per whatever millimeter or meter.

C_w will be again fF or pF. I am just going to write at fF per length of the wire. It is kind of characterized, the interconnect, so the wire is kind of characterized for unit length of the wire, per unit length of the wire. If I have l/N here, the overall resistance of this particular segment, segment number 1 will be $R_w \frac{l}{N}$, the capacitance for this segment will be C_w l/N.

Similarly, for this segment number 2, I will have $R_w \frac{l}{N}$, $C_w \frac{l}{N}$. For the third segment I will have $R_w \frac{l}{N}$, $C_w \frac{l}{N}$ and so on till it reaches the Nth segment, where we will have $R_w \frac{l}{N}$ and then $C_w \frac{l}{N}$. The first thing is the repeaters have been added here at each of the N - 1 segment, so that my overall signal strength while it is the signal starts from A and while it reaches to the B, we will have a good signal strength, not an attenuated anymore.

Remember that this particular repeater it is nothing but in the form of an inverter. If I consider the inverter here at each of the repeater points N - 1 points, I will have a V_{dd} rail or I will have a ground rail. What it implies is at the output of this, it will be either be driven by a strong V_{dd} signal or a strong ground signal. Based on the input signal, we will get the output to be a stronger logic 1 or a stronger logic 0.

In that sense I will have a strong signal here while it reaches to the point B. If I have even number of inverters here as a repeater; then in that case my logic whatever is there at the point A will be same as that will be reached at point B. If I have odd number of inverters, of course the logic will be complementary at point B with respect to the A. There is one another way of having the repeater is each of this point, instead of an inverter, it could be a pair of inverters

An inverter very very close to another inverter, will have a pair of inverters. In that case at each of this repeater points, I will have a pair of inverters and at point B I will get non complementary logic. There are different styles of designing the repeater, but the overall aspect of the repeater is making sure that we will have a non-attenuated signal at point B at the other end of the wire.

Coming back to this particular the segment. If I have created N segments, the overall or rather the delay of one particular segment. I am going to write it as a delay of one particular segment as an RC delay, it will be nothing but,

$$\text{delay}_1 = R_w C_w \left(\frac{l}{N}\right)^2$$

The overall delay for the wire which is of N segments will be nothing but,

$$\text{delay}_1 = \sum R_w C_w \left(\frac{l}{N}\right)^2$$

$$\text{delay}_{\text{wire}} = N R_w C_w \left(\frac{l}{N}\right)^2$$

$$\text{delay}_1 = R_w C_w l^2 / N$$

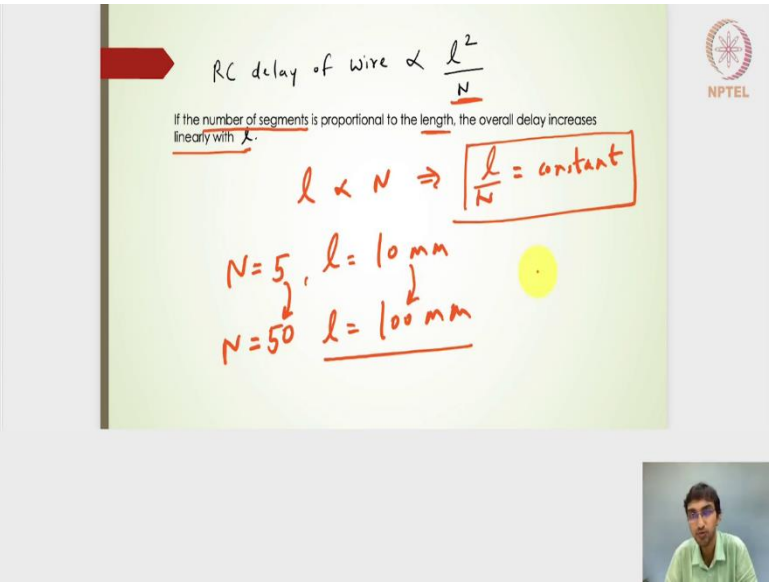
Now, this is very important, it is still even if I segment into N segments and add the in the repeaters here, if I increase the length of the wire that RC delay of the wire is still proportional to the square of the length of the wire, so it is not going to get eliminated.

But if somehow, I make this l/N as a constant, what it means is, if I make this as a constant. If I, have based on the length of the wire, if I can proportionally add the number of repeaters, if I proportionally divide the length of the wire into N segments, if the l is 100 millimeter.

Then with respect to 1 millimeter whatever the number of segments I have created, it remains linearly proportional. The number of segments I will create for 100 mm will remain linearly proportional with respect to the number of segments I have created for a long length of wire of 1 mm.

In that case if l/N is a constant or it is directly proportional to the number of segments, then I can actually eliminate, the RC delay will now be directly proportional to the length of the wire.

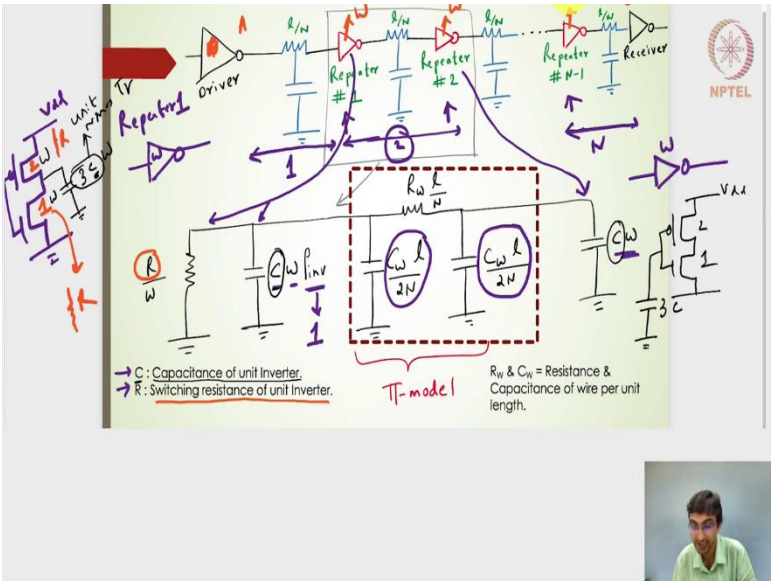
(Refer Slide Time: 14:42)



Coming to this particular slide. What I mean is, RC delay as we have seen is directly proportional to the $\frac{l^2}{N}$, but if somehow the number of segments is made proportional to the length of the wire, the overall delay increases linearly with l, which is a good factor compared to when we had an $\frac{l^2}{N}$ or without even making the segments, it was directly proportional to the l^2. In that sense if I somehow make $l \propto N$, that means l/N is some kind of a constant. Then my RC delay of the overall wire will be linearly proportional to the l of the wire, the length of the wire. What it really means is, if suppose the number of segments I have created is, let us say N = 5 for the l = 10 mm.

Tomorrow if I want to design in some other portion of the of the chip a 100 mm wire, then number of segments I will have to create for this particular length of the wire will now be 50. It increases proportionately as the length increases and in that case my RC delay of the wire will now be directly proportional to the length of the wire, it will have a linear relationship with the length dimensions of the wire and not the square of the length of the wire. Hope this is clear.

(Refer Slide Time: 16:20)



So, if we understand that, then I think we will now try to arrive at that particular relationship between the length of the wire and then the number of segments. In the last slide we had seen a segment of the wire and then I had made N segment. Here what we have done is, at each of the segments not only the segments are created, but also a repeater is added.

Here we have created this is point A and then let us say this is point B and at each of these segments wherever the segments are created, there is a repeater been added, this will be my segment 1. Let us say that this is my segment 1 and this is the segment 2, then the segment 3, segment 4, segment 5, N - 1, this is the segment N.

I have created N segments and not only created the N segments, but also I have added the repeaters here repeater 1, repeater 2, repeater 3, repeater N minus 1. Now, if I add the repeaters, then I will have each of the segments of length l/N. That is what I have written here l/N, l/N and l/N. If I look into this particular aspect of the wire or rather the second segment of the wire, where we have the l/N length of the wire and on each side I have a repeater number 1 and then the other side I have a repeater number 2.

In this particular case I have drawn an inverter. I have just taken a simple example of the digital circuit and instead of a repeater or rather in the form of a repeater I have added the inverters. I have an inverter 1 and inverter 2 and then in between that I have the one

segment of the wire, the length of the segment of the wire is l/N and each wire can be represented in the form of a pi model.

$R_w \frac{l}{N}$will be the resistance of that particular segment, the overall capacitance is $C_w \frac{l}{N}$, but in the π model we know that half of that particular capacitance will go in one side of the resistance and on the other side of the resistance, other half of the capacitance will go. That is why the capacitance of $C_w \frac{l}{2N}$ and $C_w \frac{l}{2N}$ is there on the two of the sides of the resistance.

The repeaters which is nothing but an inverter here and it is switching resistance and then the parasitic capacitance is drawn here. The parasitic capacitance is I have stated it as C, based on the width of the inverter, we will have C_w. This is a normalized parasitic factor written as a rho inverter. Up till now we had seen the parasitic factor of an inverter to be nothing but 1.

Then in the immediate future slides we will see whether it can be less than that of 1, then it will be good, because my overall capacitance will be less, which will help in reducing the overall delay. This is the switching resistance of the repeater number 1 or the inverter in this particular case. C_w turns out to be the input capacitance or the gate capacitance for the repeater number 2.

Let us say that this is the repeater number 2 what I have drawn here, if it has a width of w, then I will have a CW. If the repeater number 1, which is nothing but this is the repeater number 1 and which is nothing but an inverter number 1 and if it is of size w, then I will have $CW\rho_{inv}$. Remember that R_w and then C_w are the resistance and the capacitance of the wire per unit length of the wire.

Then C and R are nothing but the capacitance of the unit inverter and then the switching resistance of the unit inverter. Till now what we had seen for a unit inverter is. Let me draw the unit inverter and this is my unit inverter with the V_{dd} rail and then the ground rail. Let me take up the unit inverter as 2:1. In my 2:1, I always used to have a 3C capacitance, because 2C capacitance coming from the PMOS side and 1C capacitance coming from the NMOS side.

Now, in this particular case, this C is actually nothing but 3C capacitance, where C in this particular case was nothing but the unit NMOS transistor. That is what we used to take

right, that particular C value is nothing but a unit NMOS transistor, whereas here this particular C is nothing but this 3C of the unit inverter. Just for your own understanding, this C is actually nothing, but 3C, if this C is nothing but unit NMOS transistor.

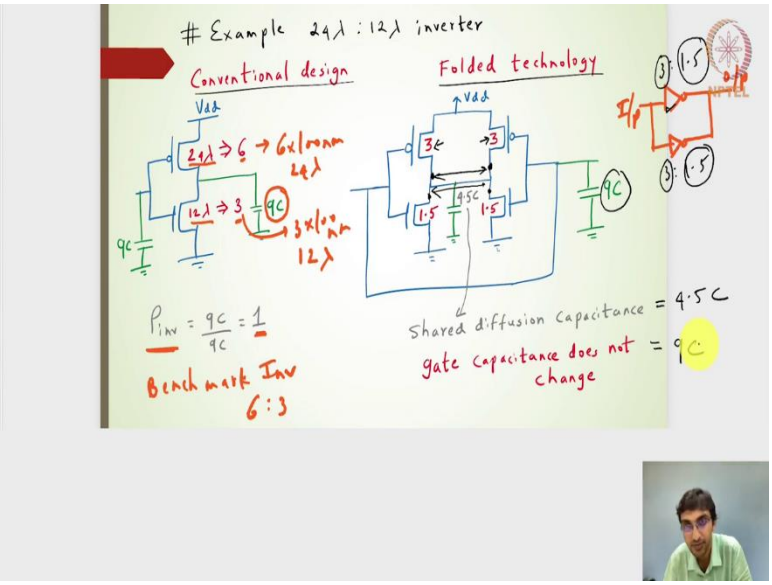
But I have consider this C as nothing but the capacitance of a unit inverter, that is all the difference. If it is scaled by w times, then I will have 3CW. Then the parasitic factor if it is 1, it will be CW, if it is something else, it will be $C_w\rho_{inv}$ or the parasitic inverter. Even the gate capacitance here, the same inverter I can draw here, 2:1 and the PMOS and the NMOS.

Then the gate capacitance should have been 3C, assuming the C is nothing but the gate capacitance of a unit NMOS transistor, but this C is nothing but of the unit inverter. This C includes 3 times the unit NMOS transistors capacitance. Similarly here this C includes 3 times the unit NMOS capacitance, but this C I have just written it down as capacitance of the unit inverter.

Note that particular difference with respect to our previous lectures. The switching resistance of the unit inverter will be nothing but let me rewrite this the switching resistance will be nothing but R. That is what I have stated here and then if it is scaled by w times, I will have $\frac{R}{w}$. That is what I have written switching resistance of the unit inverter.

Hope the R_w, C_w, R and C values are clear at this particular point of time. Remember that we have a driver with the scale of w or rather not the driver and the repeater with a scale of w and repeaters with a scale of w's, repeater N minus 1 with a scale of w. The R_w and C_w is nothing but the resistance and then the capacitance per unit length of the wire, hope this is clear.

(Refer Slide Time: 23:18)



Now let us have a look at the normalized parasitic factor, the normalized parasitic delay in our linear delay model. I am taking an example here of 24λ and 12λ, let us say that it is nothing but the 6:3 inverter. The 6 inverter represents 6x100nm = 600nm and 3 represents 3x100nm.

Each of the 100nm is nothing but 4λ. Normally the unit NMOS transistor has a minimum of 4λ. A scale of 1 represents 4λ, which is where λ = 25nm. If I consider that will be 12λ here and this will be 6 x 4λ = 24λ and that is what I have written here 24λ and then 12λ.

If I have the 6:3, then my output capacitance is actually the parasitic capacitance is 9C. The benchmark inverter for this particular case will be nothing but 6:3, because I need the same output current whether it is discharging current or the charging current it will have a capacitance of 9C. The benchmark inverter will be 6:3. My normalized parasitic factor will be,

$$\rho_{inv} = \frac{9C}{9C} = 1$$

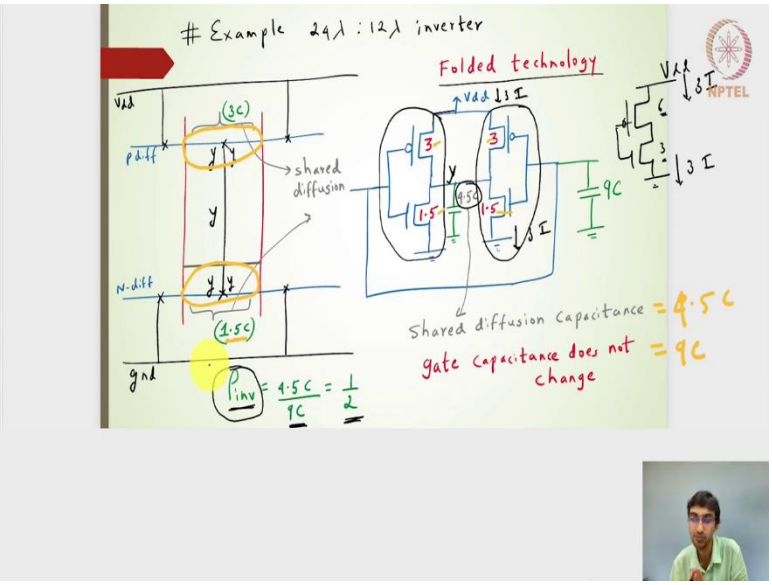
Now, if I look into the folded technology, it is kind of we have this as a conventional design and then this is a folded technology. What we have is basically two inverters, where the inputs are connected and where the outputs are also connected. I have an input here and I have an output here. The size of this particular inverter will be nothing but, if it is 6:3 here, I will make this size as 3:1.5 and 3:1.5.

Put together this 3 + 3 = 6 and then 1.5 and 1.5 on the NMOS side will be nothing but 3. Let us say that if I make this 3 and 1.5, that means 3 and 1.5 on the PMOS and NMOS width and then 3 and 1.5 on another inverters PMOS and NMOS width. This is PMOS and then this is also PMOS and then NMOS and NMOS has 1.5, my input capacitance will still be 9C.

But now the output capacitance I have written it as 4.5C, because the diffusions here in which are in parallel are kind of merged or shared now and then the diffusions here on the NMOS side and then on the NMOS side are also shared and they are merged. My overall capacitance here because it is a emerged or they shared the diffusion, I will get 3C from this particular diffusions.

Then I will get 1.5C from this particular diffusion. I will have the shared diffusion capacitance as 4.5C instead of a 9C, the gate capacitance does not change, it will still be 9C.

(Refer Slide Time: 26:45)



Moving forward, this is what I have drawn and the stick diagram here just to ensure that to give you a more detailed representation of the diffusions.

This is the P diffusion line, this is the N diffusion line, I have a V_{dd} rail, I have a ground rail. If I consider this particular portion of the inverter, the left side of the inverter, I have drawn it on the left side in the stick diagram also. I have a polysilicon gate the red line,

which divides the p diffusion into two diffusion pockets and one diffusion goes to the V_{dd}, because this is what goes to the V_{dd}, the other diffusion is my output line y.

On the N diffusion side, I will have the input, this is the polysilicon gate input, on one side it goes to the output line. This is my output line and on the other side it goes to the ground. This particular diffusion goes to the ground. If I consider this the right side inverter 3:1.5, I will draw the similar mirror image of this particular inverter in the stick diagram, I will just draw it and the mirror image.

This on one diffusion it will go to the V_{dd}, N diffusion line one diffusion goes to the ground, on the other side it is nothing but the y output, on the other side it is nothing but the y output. What it really means is, this particular part of the diffusion is actually shared diffusion, it is a shared contact; because there is a output line which gets connected from P diffusion line to the N diffusion line, anyone in the N diffusion line this is actually shared.

If this is shared, the diffusion widths will be 1.5 and 3C only and not 6C and 3C. My overall capacitance in this case for the shared diffusion capacitance will not be 9C, will be 4.5C, the gate capacitance will still be 9C, because the gate capacitance is nothing but accumulation of all the capacitance at the gate side, which will be 3, which will be 1.5, which will be 3 here and 1.5, turns out to be 9C.

The normalized inverter in this case will be nothing but the output capacitance 4.5C and then the divided by the benchmark inverter in terms of the conventional design, the benchmark inverter will still be 6:3, if I consider the charging current and then the discharging current that is going to the that is coming from the V_{dd} rail and then going to the ground rail, it will still be the benchmark inverter will still be 6:3.

Because at the V_{dd} the charging point and in the ground rail the discharging point, the current that goes to the ground rail for while it is coming down will still be 3I and while it is charging from the V_{dd} rail it needs to be still be 3I. As compared to the current is actually 3I here and while it is discharging to the ground, the current is actually 3I. The output current while it is discharging or while it is charging for 4.5C, it will still be 3I.

For defining my benchmark inverter, it has to be a conventional benchmark inverter should be 6:3. My normalized inverter in this particular case will be,

$$\rho_{inv} = \frac{4.5C}{9C} = \frac{1}{2}$$

Hope this is clear. Now, this forms a very important aspect, this particular ρ_{inv} will be included in our repeaters especially in the delay part. To estimate the delay as well as to estimate the energy, we will see this particular ρ_{inv} which is nothing but it could be 0.5 or it could be 1 depending on what kind of the design technology we will be using.