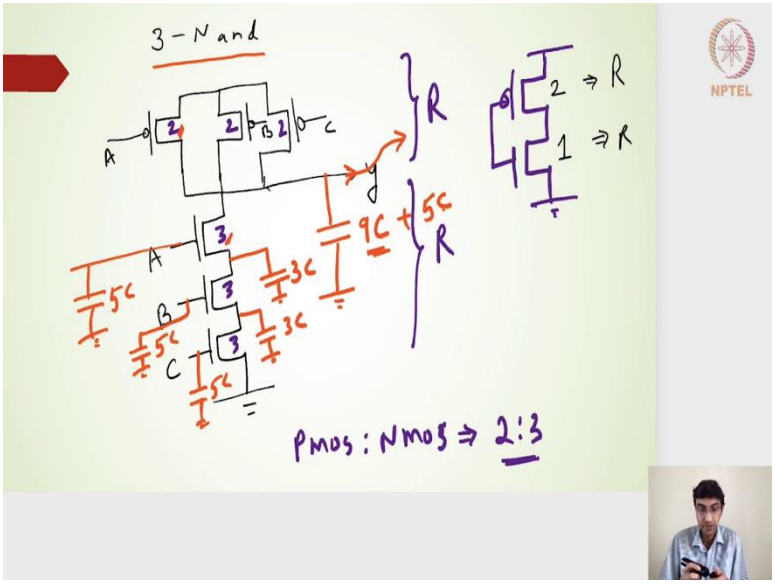
Design and Analysis of VLSI Subsystems
Dr. Madhav Rao
Department of Electronics and Communication Engineering
The International Institute of Information Technology, Bangalore

Lecture - 28
Characterizing Delay of NOR gate

Hello students, welcome to this lecture on exploring the delay of the NAND and the NOR gates. Let us proceed further, I think its an empty slide, what it means is I need to summarize the 3 input NAND gates and how did we estimate the delay of the 3 input NAND gate.

(Refer Slide Time: 00:40)



The 3 input NAND gate, I am going to summarize this again. The pull up circuit for the 3 input NAND gate to be in parallel, all the PMOS transistors will be in parallel. This is my PMOS transistors shown with a bubble sign, bubbled input that represents our PMOS transistors.

Then 3 of the NMOS transistors will be in series in the 3 input NAND gate and this will be output volt. Just to make the inputs A and A here and then the B input and then the C input and then similarly I am going to represent the B and the C inputs on the pull down side.

What should be the sizing? You know as per our our understanding we need the rising resistance equal to that of the rising resistance of the unit inverter. The unit inverter I am going to draw the PMOS and NMOS are tied inputs and then this will be my V_{dd} and then this will be ground and for the unit inverter we need a size of 2:1. So, as to give us the resistance equivalent to R.

In the similar way if I want the rising resistance and then the falling resistance is equal to R here and then the R here. The sizing of this particular 3 input NAND gate should be 3, 3 and 3 and then here, in the worst-case condition one of the PMOS transistors will be on and the other two will be off and making this as the size as 2 and 2 and 2.

What we are seeing is a PMOS and NMOS size ratio for a 3 input NAND gate, NMOS is nothing but 2:3. This is what we are going to represent from here on 2:3 for a 3 input NAND gate and 2:1 for an inverter gate. Also we had estimated the delay by utilizing the capacitance or by expressing the capacitance at the output node and at the inter diffusion nodes as well.

At the output side I am going to represent an equivalent or an accumulated capacitances. Coming from the diffusions of the transistor A, B and C from the PMOS side and then from the transistor A on the NMOS side. Overall, I will get $2 + 2 + 2 + 3 = 9C$ at the output node.

The inter node or diffusion capacitance will give me a capacitance of 3 here, because of the merged uncontacted diffusion nodes and then here also we will get the capacitance of 3, this is about the 3 input NAND gate as a standalone gates, I am going to see a capacitance of 9C.

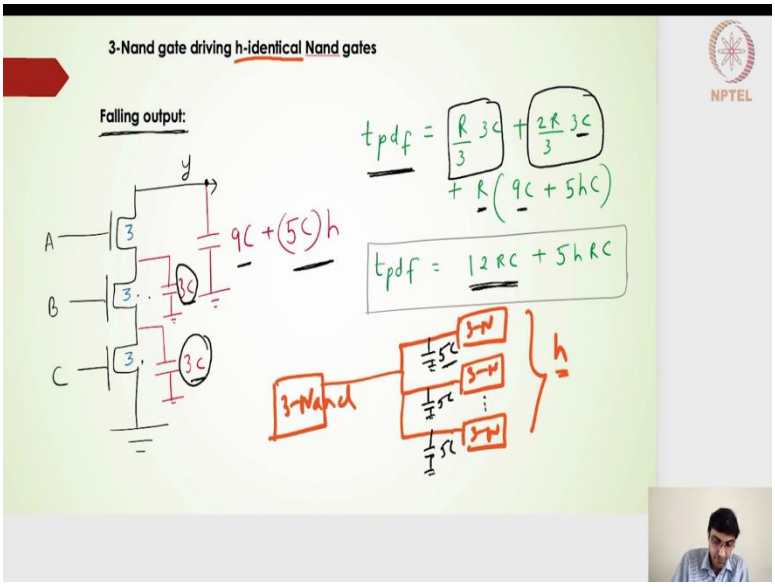
Now if I make this output node fed to another 3 input NAND gates or multiple 3 input NAND gates. What I am visualizing is one the first stage is a 3 input NAND gate. The second stage the output of the first stage is fed to the second stage inputs, one such 3 input NAND gate or multiple such 3 input NAND gates in the second stage.

Remember that when I connect this to the next stage the input capacitance of the next stage will be considered as a load capacitance for the first stage. The load capacitance for the first stage will be nothing but its own parasitic capacitance of 9C + input capacitance coming from the next stage.

If I am connecting this 3 input NAND gate to another 3 input NAND gate of the similar dimensions, the input of the second stage 3 input NAND gate will be nothing but 3 + 2 because that is the input, each of these inputs sees a capacitance of 5C.

The B also gets a capacitance of 5C and C also sees a capacitance of 5C. If I am connecting this output to the next stage of a 3 input NAND gate. Again, I should get additional 5C capacitance are considered in the load capacitance for the first stage, hope this is clear and moving on.

(Refer Slide Time: 05:47)



That is what we have in this particular slide, I am talking about a 3 input NAND gate driving h identical NAND gates and what do you mean by the h identical, that means the size of the second stage NAND gates are same. It is basically something like this, the first stage you have a 3 input NAND gate and which is then fed to h such 3 input NAND gate. This the number of the 3 input NAND gate is nothing but h.

In that case each of this 3 input NAND gate is going to give me a capacitance of 5C. Each of this 3 input NAND gate is going to give capacitance of 5C, capacitance of 5C, h number of times will give me 5Ch, 9C is its own parasitic capacitance or the diffusion capacitance or also we can call it as a depletion capacitance.

9C + (5C)h will be now we considered across the output node. Its own inter node diffusion capacitance or also called as merged uncontacted diffusion capacitance is 3C and then 3C

here. I have drawn it only for the pull down circuit because I am considering only the falling output at this particular stage and I want to estimate what is the propagation delay for the falling output.

This is what the propagation delay falling output turns out to be nothing but, you will use now one can use the Elmore delay method and then estimate what is the propagation delay for the falling output. This particular transistor C which is closest to the source is going to contribute to a switching resistance of R/3.

The B transistor is going to contribute R/3 and then A transistor is going to contribute resistance of R/3.

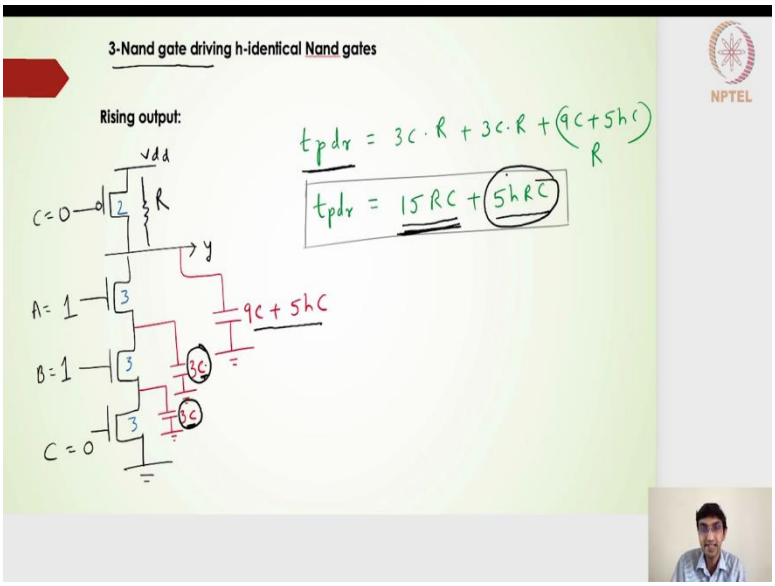
$$t_{pdf} = \frac{R}{3}3C + \frac{2R}{3}3C + R(9C + 5hC)$$

We will effectively get the propagation delay falling for a 3 input NAND gate driving h identical NAND gates will be nothing but,

$$t_{pdf} = 12RC + 5hRC$$

Earlier when it was not connected to the h identical NAND gates we got the propagation delay falling as 12RC, if it is now connected to h number of 3 input NAND gates, each of the NAND gates incorporates a capacitance of 5C and that is why we are saying 5hRC, hope this is clear.

(Refer Slide Time: 09:12)



Moving forward what is the propagation delay for the rising output for the same configuration. The 3 input NAND gate driving the h identical NAND gates for the rising output. Again, we will consider the worst case condition is one of the transistor is on and the other two will be off.

Nevertheless, on top of it, we have on the pull down side the two transistors A and B which are away from the ground because C is closer to the ground rail. The two transistors which are close which are away from the ground and then closer to the output node are on. The worst-case condition I will have this 3C capacitance and these 3C capacitance also to be considered in our delay estimation. Thereby the increasing our delay and thereby giving the worst-case delay.

In such a scenario when these two transistors are on the PMOS on the pull down side on the NMOS transistor. That means, on the PMOS side the A and B transistors will be off and C is only on, that means, C=0 here, that means, that the C=0 here represents that the C transistors on the NMOS side is off.

Now I am left with 3 such capacitance, one is a 9C + 5hC because this output node is connected to h, such 3 input NAND gate on the second stage I am having 9C + 5hC. Then I am having this 3C and then this 3C capacitance. Using again the Elmore delay method I should be able to find out the propagation delay rising,

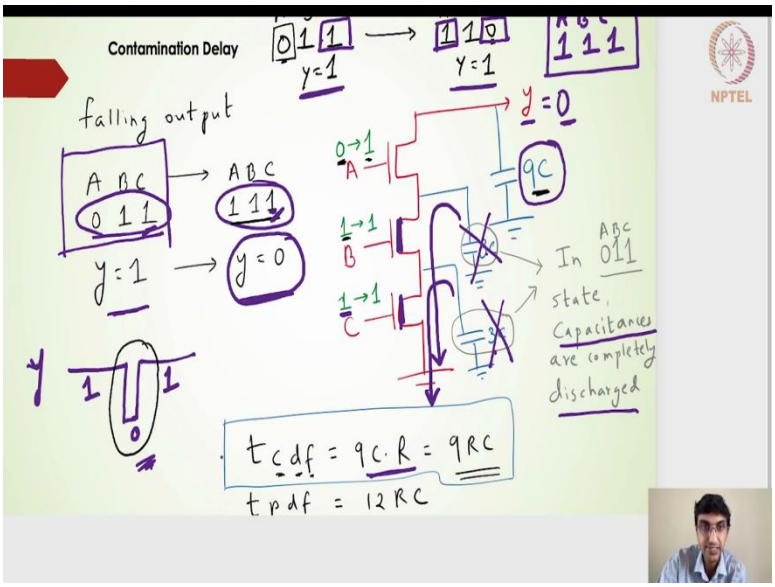
$$t_{pdr} = 3CR + 3CR + (9C + 5hC)R$$

$$t_{pdr} = 15RC + 5hRC$$

When we considered the rising delay for a 3 input NAND gate without the h identical NAND gates, without cascading to the next stage, we got 15RC as a rising delay here because of the 5hC parameter, we are getting 5hCR for the rising delay that is an additional coming from the next stage capacitance.

That is going to add to the delay, hope this is clear. What we have seen until now is the falling delay and then the rising delay for the two stages, where the first stage is a 3 input NAND gate and then the second stage is the h identical NAND gates.

(Refer Slide Time: 11:50)



In this particular slide we are going to look at the contamination delay for the 3 input NAND gate. Again, we will look into the falling output, the contamination delay for the falling and the contamination delay for the rising output. Let us take a look at the falling output. What we really want is the input, let us say that it starts from 0 1 1 and then changes to 1 1 0 for a 3 input NAND gate.

If I am going to write it as A B C here and then this is what the A B C represents. The output for both the states for a 3 input NAND gate this is 0, that means the PMOS transistor is on and then the output will be charged to V_{dd}. In this case also the output should be charged to V_{dd} because one of the transistor is on.

On the NMOS side one of the transistor is off, that means that the output node cannot discharge, it can only charge to V_{dd}. The output should be y here, y =1 and then y=1 here. But for some reason the input when it does a transition for example, if I take this A input it does the transition from 0 to 1, B does not change at all.

There is no transition for B and there is a transition for C, 1 to 0. Let us say that for whatever reason you know not all the inputs does the transition at the same time. It is quite possible that A will do the transition from 0 to 1 earlier than that of the C. That means, at time t=0ns or 10ps we will have the input A doing the transition from 0 to 1.

But after 10 more picoseconds, the C will do the transition from 1 to 0, in between that 10ps after the A has done the transition the input parameters, the input values will be actually be A has done the transition from 0 to 1, B stays at 1, C has not done the transition, A B C is actually 1.

What it implies is the A B C value now which makes the output go down, because now all the NMOS transistors are on and the PMOS transistor all of them are off, that means that the output y which should have been retained a value of 1 is now completely discharged to 0, there is a path for the discharging.

In that case from 0 1 1 to 1 1 1 here in this particular case, the output was 1 and then it is doing the transition to 0 which is a contaminated output, this is something we do not want, what we wanted was a 0 1 1 input. Going to the next state of 1 1 0, that means the output should not have been changed it should have retained y=1 for both the states.

But because the input transition has arrived late for one of the input transitions has arrived late and one of them has arrived earlier. In that case we do not have that control and we will get a small output glitch from 1 to 0. Then after that once we have the C input doing the transition from 1 to 0 this y will go from 0 to 1.

But there is a small gauge figure, that means if I draw the y output profile, I will actually get 1 here and then there will be a small glitch and then it will go back to 1, this is 0 and then this is 1. This is a contaminated output and that is why we are calling it as a contaminated delay and especially because its a falling output we will call it as the falling contaminated C_{df} contamination delay for the falling output.

If I look at this schematic here and its associated capacitance we should be able to derive the contamination delay for the falling output. The output capacitance is nothing but, 9C and inter node diffusion capacitance is 3C and 3C. If I have A doing the translation from 0 to 1, B is anyway is 1 to 1 and then C is not nothing but 1 to 1. If I am taking this middle state transition A is changing from 0 to 1.

Earlier it was 0 and now it is changing to 1, that means this associated capacitance here on the 9C at the output side 9C. There was no channel path for the 9C capacitance to discharge to 0. Earlier when A was 0, there was no channel formed here, that means that there was no discharging path for the 9C capacitance.

The 9C capacitance was earlier charged to V_{dd} value and the moment A goes to 1 it should start discharging, hope that is clear. B was earlier 1 and now also it is 1 in this particular intermediate state of 1 1 1. That means this 3C capacitance earlier when the B was 1 here this particular 3C capacitance would have completely discharged.

C was also 1, in this case C is 1 here and then C is 1 here, so see the earlier one. There was the channel formed in this particular C transistor and thereby we get this particular channel is also formed. The discharging path for both this capacitance was earlier available and it would have completely discharged to 0.

What it means is in the earlier state of 0 1 1 the capacitances are completely discharged, which capacitances, the 3C capacitance which are nothing but finding a discharged path through the on transistor are nothing but B and C on transistors.

The contamination delay is nothing but all the delay associated with the 9C which is the output node capacitance. The 9C multiplied by this R/3, this R/3 and this R/3 will form the discharge path or the switching resistance and we will get,

$$t_{cdf} = 9C \times R = 9RC$$

We will not consider this 3C here and then we will not consider this 3C here because it is completely discharged to 0. The moment the output at this particular state itself 0 1 1 both these capacitances 3C and 3C capacitances are discharged to 0.

The only capacitance that has to discharge to 0 to give us the output of 0 here will be this 9C capacitance. Thereby it is called as I mean, so that is why the delay is associated with

this particular 9C capacitance and not with this 3C and 3C. But since this output 0 is a contaminated output that is why we call it as a contamination delay for the following, turns out to be 9RC.

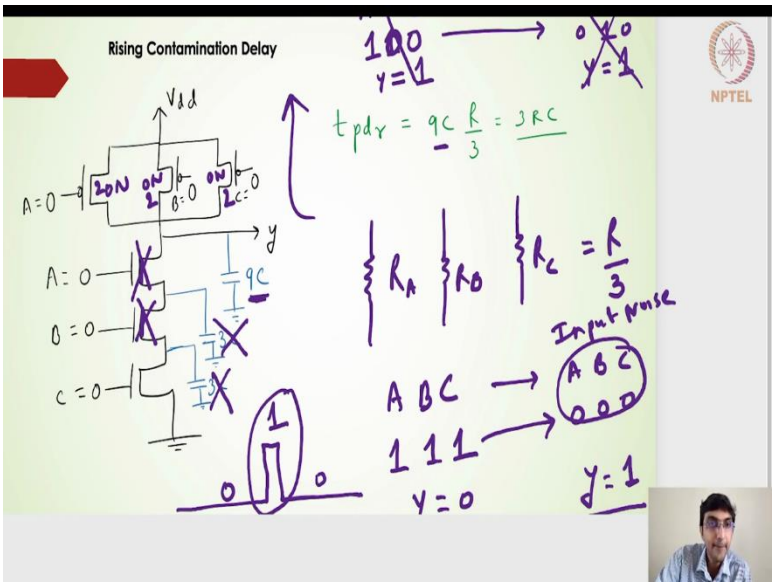
If at all that this output was cascaded to the next stage with 5hC capacitances seen as a load capacitances, because we have the h number or 3 input NAND gates. In the next stage then this particular value would have been 9RC + 5hRC. Remember just to compare with that of the propagation delay falling for the 3 input NAND gate it is nothing but 12RC.

$$t_{pdf} = 12RC$$

If I am comparing the contamination delay and the propagation delay, the propagation delay should always be higher and contamination delay should always be the lower value. I think this contamination delay is kind of very very it is a one of the design parameters that we need to consider especially while we designing the sequential elements. When once we have the flip flops or the latches and in between the combinational circuits and then followed by one more sequential elements.

We have to consider the contamination delay, so that the hold time where violation does not happen and the hold time violation is characterized for a particular sequential element. When we talk about the sequential elements we will once again take this particular or involve this particular parameter the contamination delay and that is why all of you should be able to understand the contamination delay and that is why it is there in this particular lecture.

(Refer Slide Time: 20:57)



Moving forward, what should be the rising contamination delay. The rising contamination delay is a very straightforward, if all of them I want a rising output here. For always for the propagation delay rising, we are always considered the worst-case where one of the transistor is on, the other two will be off.

For the rising one, the best-case delay is nothing but when all the 3 transistors are on, transistor A is on, transistor B is on, transistor C is on. If all the 3 transistors are on the equivalent switching resistance seen from the PMOS side is nothing but the switching resistance of A, in parallel with switching resistance of B, in parallel with that of the switching resistance of the C.

The equivalent resistance will be nothing but R/3 because each of this gates are of the size 2, this is the size 2 this is the size 2 each one of them will g offer a resistance of R each and then 3 of them in parallel, we get R/3. The output node capacitance is 9C and if this 0 this is 0 and this is 0, we will not consider this 3C capacitance because there is no charging path; there is no path here, there is no path here, I mean for this 3C capacitance to charge.

This particular path is not there, this particular path is also not there. The only thing which C is the charging path is the 9C capacitance.

$$t_{pdr} = 9C\frac{R}{3} = 3RC$$

This will be the rising contamination delay. Now in our previous example of the falling contamination delay, we had taken two states right 0 1 1 to 1 1 0 and intermediate state was 1 1 1 and that is where we said that it is not contaminated, it gives us a contaminated output. Here we have not considered that what I am saying is, if my state for whatever reason if I had 1 1 0 as my initial state or rather 1 0 0 as my initial state.

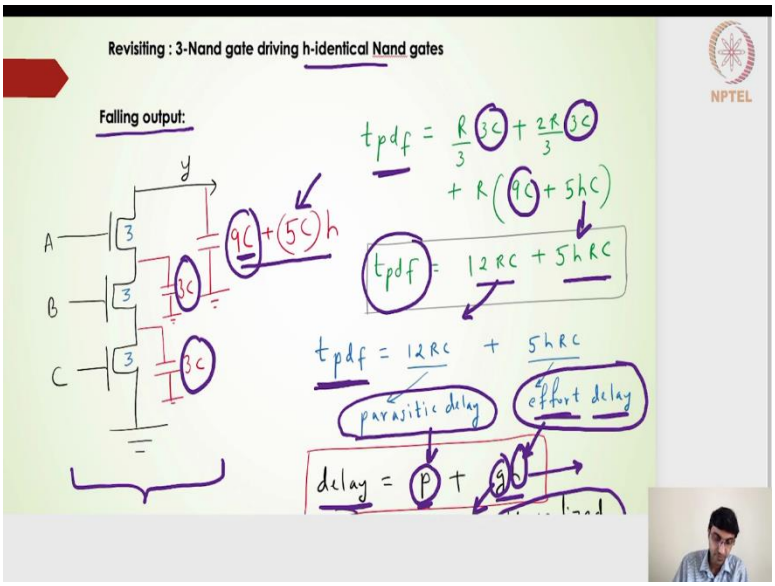
The output were 1 0 0 as an input, if this is the input A B C as 1 0 0. The output for this A B C of 1 0 0 will be nothing but 1. Let us say it changes to A B C of 0 1 0, the output is still one, I think know this is possibly this is a kind of a bad example. Let us say that the output is you know the input is A B C and it is 1 1 1, so I will get the output to be 0 and it should have been straight to 1 1 1, but whatever reason at the input side of A B and C if we get a lot of noise signals. Then we can have this state of A B C as 0 0 0, where the output will now become 1, this is all induced due to the input noise signal, but then the noise will have you know it could be a very high frequency or something. We will see a kind of a glitch here the input doing the transition from 1 1 1 to 0 0 0 in a very very small time duration and then get back to 1 1 1.

In that case my output will be 0 throughout, but then there will be a glitch of 1 and then it will come down to 0. The contamination delay could be due to the problems at the input side and the input transition does not you know for if I have a multiple inputs the inputs does not do the transition at the same time, that is one reason.

The other reason could be the input noise signal which can change the logic level at the input side and thereby the output state also changes and it all it all happens momentarily, it is not there as a steady state. You will see the glitch and then for the next steady state input you should get an out steady state output.

All this are momentary output changes and that is where we are considering the contamination delay, because all these glitches are contaminated outputs, hope you understood this.

(Refer Slide Time: 25:56)



Here we are revisiting the 3 input NAND gate driving the h identical NAND gates, this is something we had already seen. Sizes of the 3 and 3 are written and on the PMOS side the size is nothing but 2 for the each of the PMOS transistors. I am considering only the falling output that is why I have taken only the pull down circuit, that means only the NMOS transistors is in the picture. We have drawn the output node capacitance and then this is coming from the next stage. 5Ch because h identical NAND gates are there in the second stage.

This output is actually fed to the h identical NAND gates and the 9C is coming from its own parasitic capacitance. Overall propagation delay falling is

$$t_{pdf} = 12RC + 5hRC$$

Which is something we had seen earlier in one of the previous slides. This 12RC is actually coming from this 9C, its actually coming from this 3C. Then this 5hC is going to give me 5hRC, which is coming from the next state input and then this 12RC is its own parasitic. All these capacitances 9C, 3C and 3C are its own diffusion or parasitic capacitances. That is why this particular component if I actually segment it to the two components 12RC and 5hRC.

First component is the parasitic delay, because it is coming from the parasitic capacitances. The second component which is coming from the second stage, the second stage input

capacitances is we are calling it as an effort delay. This is an additional effort in terms of delay, this particular the first stage circuit is doing with respect to the for the next stage.

It could be the next, it could be no circuit at all. In that case there is no effort delay. But if I have in the second stage there is some circuit, I will have to put some effort in the delay. My performance or the delay will be larger, that is why this is called as an effort delay.

The first stage is putting an effort for the second stage, this could also be represented as,

$$\text{delay} = p + gh$$

What it really means is, the parasitic delay component and then the effort delay called as gh. Each of this is now a value or an expression, p is parasitic delay, g is something and then h is something we have also called earlier as an electrical effort or a fan out and then g we will call it as a logical effort.

All these are the parameters for a design circuit could be a 3 input NAND gate, 2 input NAND gate, 3 input nor gate inverter and then anyone the combinational circuits and I have stated has the normalized. Generally, when I get the delay parameters for a 3 input NAND gate or any other circuit, I will get the propagation delay falling or rising.

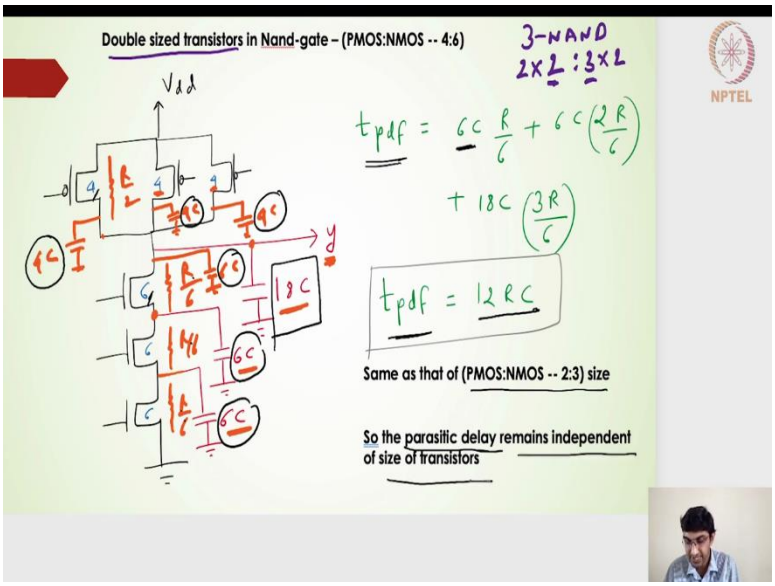
Generally, we normalize it with respect to the inverters delay. What it means is I need to always benchmark my combinational circuits or whatever the 2 input NAND gates or the 3 input combinatorial gates with respect to the inverter, the inverter 2:1 inverter which is a unit inverter.

Now according to us we know the delay parameters, it will be nothing but 3RC. If RC is 10KΩ and 0.1fF respectively, we know that this will be 3ps. We will always benchmark any of the combinational circuits or in this case it 3 input NAND gate and its delay with respect to the 3ps.

In that case what we will do is this particular propagation delay falling the absolute value of the propagation delay falling for the 3 input NAND gate, we will try to normalize it. Normalizing with respect to the 2:1 inverter. What we will do is for normalizing if we want a normalized delay component it will be nothing but this particular expression divided by 3RC.

Finally, we should be able to get the normalized component is nothing but (4 + 5)/ 3, I will come to that later. Hope you have understood this. The gist of this particular slide was, the delay component is actually segmented into two parts one is the parasitic component the other one is the effort delay.

(Refer Slide Time: 30:57)



Before getting into p plus dh format I know. Let us see what happens if I change the size of this particular 3 input NAND gate, uniformly scale it. If I double the size of the transistors in the 3 input NAND gate, instead if you remember the PMOS and NMOS in our earlier 3 input NAND gate right. The sizes was 2:3 that means, on the NMOS the size was 3 on the PMOS the size was 2.

In that case if I double this up, if I do 2 x 2 and then double this up on the PMOS and NMOS size, that means I will get a size of 4, 4 and 4 on the PMOS size and the size of 6, 6 and 6 on the NMOS size. If I have those my switching resistance here will be R/2 because it is now 4.

That means on the PMOS side if the width is scaled by k times it will be 2R/k. 2R/4 will be nothing but R/2, even this one will be R/2, even this one will be R/2. The switching resistance on the NMOS side will be nothing but R/6, this will be R/6 again and then this will be R/6.

However, if I scale the dimensions of the transistors especially the width, my capacitance here is going to increase. The diffusion capacitances or the parasitic capacitances. If I look into this particular capacitances it will not be 3C it will be 6C x kC.

This particular diffusion capacitance which is going to give me 6C, the diffusion capacitances here coming from this particular transistor will be 4C, this particular capacitance will be 4C. Particular capacitance will be 4C and I know this particular capacitance because it is 6 size it will be 6C.

My overall load capacitance seen at the output side for a single stage the input NAND gate will be nothing but this particular 6C. I will let me use a different marker 6C + 4C + 4C + 4C=18C, instead of a 9C capacitance especially when we are using 2:3, 3 input NAND gates we had 9C instead of that now it is 18C.

What we have deduced is if I change the size of the transistors if I increase the size the capacitance, the parasitic capacitance is going to scale up, the switching resistance is going to scale down. If I try to find out the propagation delay falling, I will use the Elmore delay method for the following, all the NMOS transistors should be on then only I will get the output node to be discharging to 0.

$$t_{pdf} = 6C\frac{R}{6} + 6C\frac{2R}{6} + 18C\frac{3R}{6} = 12RC$$

2:3 ratio for that particular 3 input NAND gate also we got the propagation delay falling as 12RC, Even for 4:6 ratio, that also will we got the propagation delay falling is 12RC. What I have done is I have not considered the next stage. I consider only this particular first stage change the size accordingly and then determine that the propagation delay falling does not change. That is what I have written here does not it is the propagation delay falling is the same as that of the PMOS and NMOS of 2:3 sides.

What we are concluding is the parasitic part of the delay which is nothing but, its associated with its own capacitances now what I had said in the earlier slide. The parasitic delay and not the effort delay, effort delay is coming from the next stage. The parasitic delay is due to its own stage or the own parasitic or the depletion or diffusion capacitances, that is not going to change that remains independent of the size of the transistors.