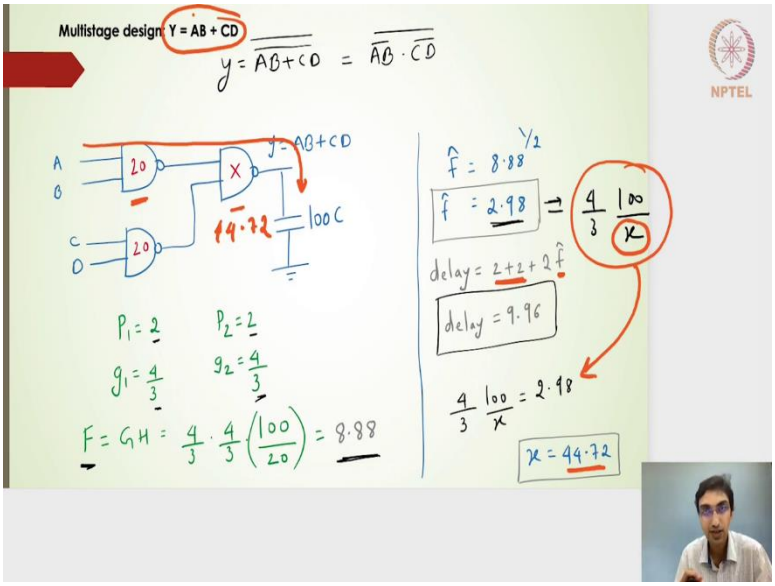
Design and Analysis of VLSI Subsystems
Dr. Madhav Rao
Department of Electronics and Communication Engineering
Indian Institute of Information Technology, Bangalore

Lecture - 39
Introduction to Combinational Circuit and asymmetric gates

Hello students, welcome to this lecture on the Combinational Circuit Families. In this particular set of lectures you will see different combinational circuit families and instead of the regular CMOS structured circuits. We try to adopt different widths making it asymmetric gates. We will adopt different family of circuits called as the passed transistors or the transmission gate circuits and then there are more towards that alright.

(Refer Slide Time: 00:53)



Let us begin with a simple example, let us try to understand different techniques of designing a combinatorial circuit. This is one such example I have taken where we want the output to be equal to nothing but,

$$Y = AB + CD$$

Where AB and CD are the inputs.

I can have this in the form of a multi stage design, multi stage means the first stage, I will have some kind of a gates and then the output of the first stage will be fed to the second

stage, that is called as a multistage design configuration. Here you can see that I have designed it using three of the NAND gates, where the two NAND gates are in the first stage and then the second one is in the second stage or rather the third NAND gate is in the second stage, Y=AB+CD it can also be rewritten as nothing but,

$$Y = \overline{\overline{AB + CD}}$$

$$Y = \overline{\overline{AB}.\overline{CD}}$$

In this particular example I have taken some specification saying that the input should see a total capacitance of 20. All the four inputs C side capacitance of 20, that means the gate size is nothing but 20 and the output I have taken as a 100C, where 1C represents the unit NMOS capacitance.

In the multistage design, what we will do is? We will follow the linear delay model. We will find out or evaluate, what is the normalized parasitic of this particular NAND gate? We will take this particular critical path, let me rewrite this let me make a mark here, this will be our critical path. In this particular critical path I need to find out the normalized parasitic for this particular gate and then this particular gate.

That is what I have written here the normalized parasitic of a 2-input NAND gate is nothing but,

$$P_1 = 2$$

Normalized parasitic of the next stage 2-input NAND gate is,

$$P_2 = 2$$

The logical effort is,

$$g_1 = \frac{4}{3}$$

$$g_2 = \frac{4}{3}$$

If I want to find out the path effort here for this particular path, it will be nothing but,

$$F = GH = \frac{4}{3}\frac{4}{3}\left(\frac{100}{20}\right) = 8.88$$

For a two stage design our path effort is 8.88, the single stage effort delay turns out to be,

$$\hat{f} = 8.88^{\frac{1}{2}}$$

$$\hat{f} = 2.98$$

Then from 2.98 I should be able to find out what should be the size of this particular gate, because I think the first stage NAND gates we anyways know what is the size here because it is given as 20 and 20.

We need to find out, what is the size of this particular gate? and then that is possible, if I can evaluate from the $\hat{f}$ which is nothing but a best individual stage effort I can equate it to,

$$\frac{4}{3}\frac{100}{x} = 2.98$$

$$x = 44.72$$

The overall delay if I have the size of 44.72 here which is nothing but the best size for a two stage design.

The minimum delay turns out to be,

$$\text{delay} = 2 + 2 + 2\hat{f}$$

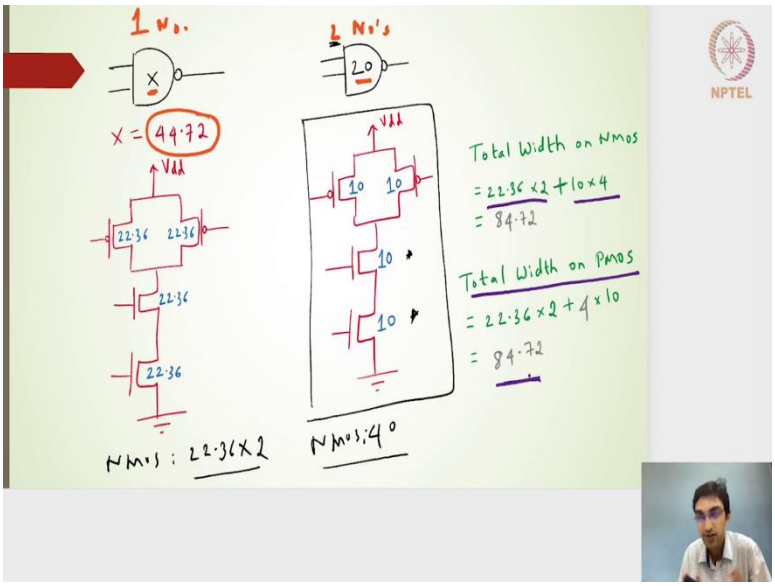
$$\text{delay} = 9.96$$

The number of the total parasitic transferred for this particular path will be nothing but the two parasitic coming from this first stage 2-input NAND gate and then two parasitic coming from the second stage 2-input NAND gate.

The overall delay is 9.96 for a multi stage design. Now, we can actually redesign the circuit Y=AB+CD.

We can redesign it as a compound gate, that means it will be a single stage and we will make conduction topology, wherein we will have a set of NMOS transistors, either in series and parallel and the replica of that or mirror replica of that the conduction complement replica of that will go into the PMOS side.

(Refer Slide Time: 06:06)



But that will be a single stage circuit, just to complete our multistage design this X = 44.72 and then in the first stage we had a size of 20 there were two such 2-input NAND gates. I am going to write it as 2 numbers here representing the 2-input NAND gate of the gate size of 20 forming the first stage.

Here we have one number because in the second stage we had only one 2-input NAND gates with a 44.72 size. If I want to find out what is the width of the NMOS and PMOS transistors we know that for a 2-input NAND gate we should have a size of 2 and 2 and 2 in the PMOS side and then 2 in the NMOS side.

We will make four parts of this, two parts will go to the PMOS side and then two parts will go to the NMOS side. That is how we have done, equal parts 22.36 and 22.36 will go into the respective NMOS and PMOS transistors. For this particular the first stage gate

size of 20, we know that at the PMOS and NMOS should get the equal parts, 10 of this will go to the PMOS and then 10 of that will go to the NMOS.

What we have is in the first stage of the circuit? We should have one more set of the four transistors, because they are 2 in numbers. In the second stage we have only one such 2-input NAND gate. If I want to find out the total area, what we are trying to find out is what is the total width? The total width on the NMOS side is nothing but 10 + 10 and this is kind we have 2 such numbers.

I will have 10 + 10 = 20 x 2 = 40 on the NMOS side, the width on the NMOS side is 40 here. On the NMOS side the width here of this particular second stage 2-input NAND gate turns out to be nothing but 22.36 x 2.

The total width on the NMOS side turns out to be,

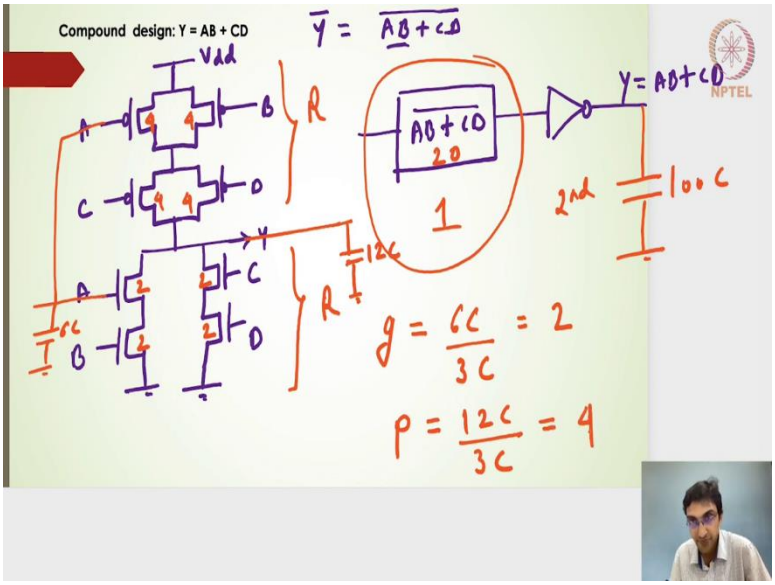
$$= 22.36 \text{ x } 2 + 10 \text{ x } 4 = 84.72$$

Then similarly we can find out the total width on the PMOS side which is nothing but,

$$= 22.36 \text{ x } 2 + 4 \text{ x } 10 = 84.72$$

What we had identified is the delay? Somewhere around 9 point something and then the overall width on the NMOS and PMOS side is 84.72.

(Refer Slide Time: 08:51)



Let me just go to the next one what we need is a compound design? I am going to draw the compound design of AB + CD. What I am going to do is, I am going to consider $\bar{y} = \overline{AB + CD}$. I am going to design a compound gate here and then add an inverter. The compound gate for this $\overline{AB + CD}$, I know that this AB will be in NMOS it will be in series.

I have now A and I have now B and this will be my output signal and I will have two legs C and D forming the transistors C and D on the NMOS side. This represents AB + CD on the NMOS side and on the PMOS side its conduction complement will come into the picture.

The conduction complement represents in the NMOS side or in the pull downside if I have two of the transistors in series that will behave, that will become parallel in the pull up side or in the PMOS side. I am going to draw two parallel transistors connected together and then similarly one more parallel transistors, this becomes my AB and this becomes my CD and then this will get connected to the output and this will be connected to the Vdd.

This is my circuit, now let us size this gates or the transistor width. That we should be able to evaluate, what should be the logical effort and then the parasitic. Because what we really want is, this particular block of AB + CD which will give me the output of AB + CD bar should go to the inverter making it the output as AB + CD.

We will take the similar dimension, the input gate size is 20 and then the output we should have it as 100C, that we can compare between this particular technique of the compound gate design with respect to the multistage designs.

Although this is a multistage design we have one stage here and then the second stage here, but still most of the computation happens in this particular part of the stage. Because it adopts all the four transistors in this particular stage, that is why we call it as more towards the compound stage design.

Let us try to assign the widths of this particular transistors. We know that to just to compare the logical effort we need the equal rising resistance and then equal falling resistance with respect to the inverter, so that my output current will be same.

If I want to have an resistance of R here. I have a size of 2 and 2 under worst case condition if A and B is 1 then I will get this particular leg to discharge having a rising resistance of

R. Similarly, if A and B is 0 and C and D will be 1 then I will have this particular leg of C and D making the discharging path for the output capacitance.

On the pull up side I will have a width of 4 and 4, 4 and 4 here. Under worst case condition either A or C, or D or A and C or D should be 0, so that these transistors will have a successful charging path, connecting the V_{dd} to that of the output node capacitance or we can have another combination of B being on, so this particular transistor being on and one of these transistors C or D should be on with respect to this particular sizing. I should be able to get the rising resistance as R. Once I have the width sizes can I find out, what is the logical effort for this particular inputs? All of them have the same widths.

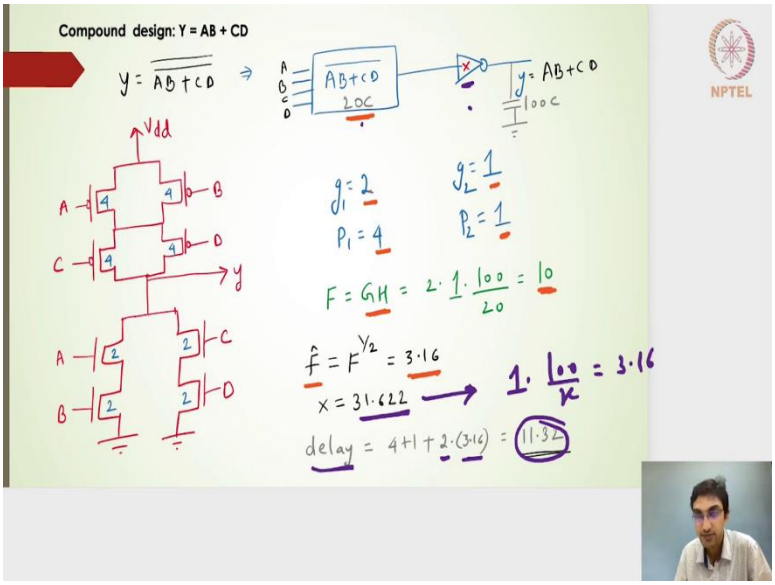
If I consider the input capacitance for one of the transistors of the inputs. I will get the capacitance to be 6C divided by our benchmark inverters which is having 2 :1 ratio, which will give me the falling resistance as R and the rising resistance as R will give me the input capacitance as 3C, it will give me the logical effort as 2.

$$g = \frac{6C}{3C} = 2$$

The parasitic again having the sizes such that we will get the same output current. This particular sizes will ensure that I will get the same rising resistance or the falling resistance with respect to the 2 :1 inverter.

If I can find out the total output node capacitance, the total output node capacitance seen at the output node, or rather the overall capacitance that is seen in the output node, if you can calculate and then take that into estimating the normalized parasitic factor. The overall output node capacitance turns out to be $2 + 2 + 4 + 4 = 12C = \frac{12C}{3C} = 4$

(Refer Slide Time: 14:42)



With this particular g and the P value here, let us try to find out the minimum delay that we can extract from this particular compound gate design technique. Here is our AB + CD with the size of 20, we know the logical effort turned out to be 2 and then the parasitic is nothing but 4, this is what we had seen in last slide.

For an inverter we have a logical effort of 1 and then the parasitic of 1. The total F the path effort in this particular case will be nothing but,

$$F = GH = 2 \,.\, 1 \,.\, \frac{100}{20} = 10$$

$$\hat{f} = F^{\frac{1}{2}} = 3.16$$

$$1 \,.\, \frac{100}{x} = 3.16$$

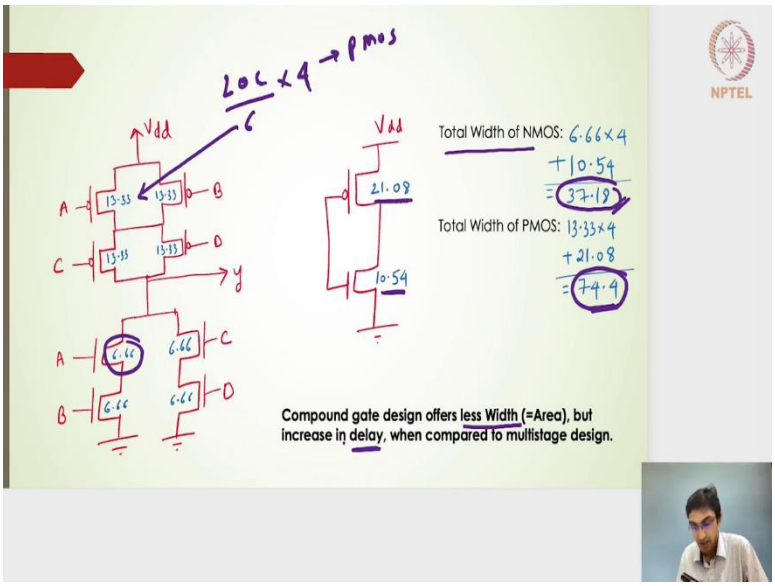
$$X = 31.622$$

If I can easily find out what is x, x turns out to be 31.622 that is the size of this particular inverter, giving me the best delay. And what is the best delay? What is the minimum delay here? It turns out to be nothing but,

$$\text{delay} = 4 + 1 + 2(3.16) = 11.32$$

The total value or the best delay turns out to be 11.32. If I go back to our previous design which was there in the slide number 2, I should be able to confirm that this particular cascaded multistage design gave us the delay of 9.96. If I go back to my present slide, this particular compound gate design gives me 11.32. The delay using the multistage design is kind of preferred, if I am considering the delay as one of the parameters. Although both of them giving the same output.

(Refer Slide Time: 17:16)



Let us also try to evaluate what is the overall area or the footprint of this particular design. I now have the sizes, what I have done is once we have evaluated, what is the X value? From the previous slide we have evaluated what is X? which is 31.622.

2:1 inverter will create three parts of 31.622 and then two parts of 30 of this three part will go to the PMOS side and then one part will go to the NMOS side. Then if I have an input capacitance of 20C or a gate size of 20 here for the $\overline{AB + CD}$.

We know that for the individual input 6 parts has to be created and then 4 parts of the 20 should go to the to the PMOS side and an 2 parts of the 20 the 6 parts should go to the NMOS side. For the 20C will make 6 parts and then 4 parts will go to the PMOS side and then the remaining 2 parts will go to the NMOS side.

This particular thing will be 13.33 is going towards the PMOS side and then the remaining two parts will go to the NMOS side. That is why we get a size of 13.33 and 6.66. With this

sizes, the widths of the transistors on both the stages, the first stage we have the compound gate and the second stage we have the inverter. The total width of the NMOS will be nothing but,

$$6.66 \text{ x } 4 + 10.54 = 37.18$$

On the PMOS side we will have,

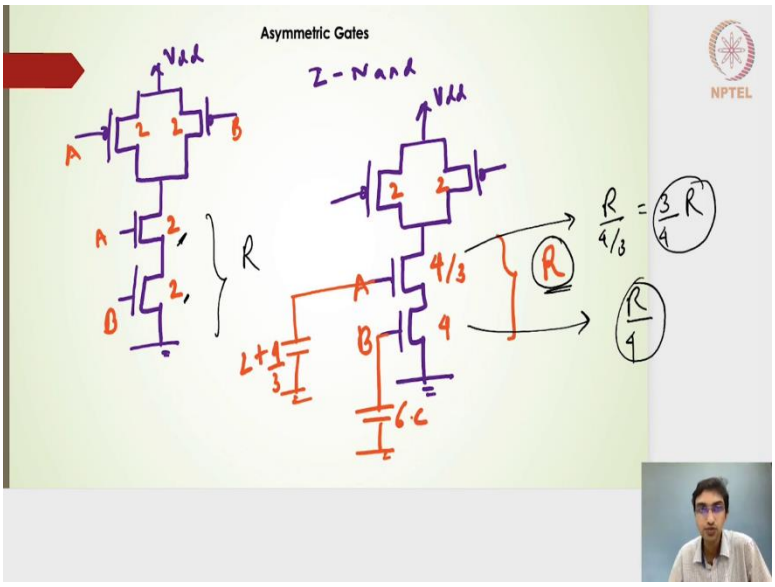
$$13.33 \text{ x } 4 + 21.08 = 74.4$$

If I consider our previous case where it was 84 something on both PMOS and NMOS. This particular size has drastically reduced 37.18 on the NMOS side, it is a drastic reduction in the total width of the NMOS side.

PMOS side also it has reduced by almost the size of 10. Compound gate design offers actually less width which will indirectly help in having a less area, having a very compact footprint of the design. But of course, we will have to compromise in the delay when compared to that of the multistage design.

This is one such perspective that is being given the output remain the same, the input gate sizes are the same, the output load capacitance being the same. But still we have two designs one giving a better delay results another one giving a better area results, hope you have understood this.

(Refer Slide Time: 20:16)



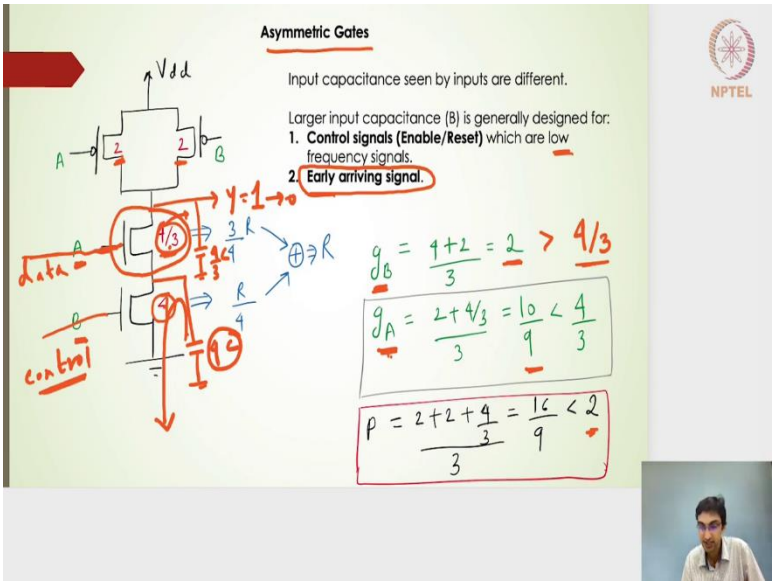
Moving on the asymmetric gates and this is an empty slide, that means that I have to draw a diagram here. Let us start with an 2-input NAND gate, what we have seen is a 2-input NAND gate with a same sizes on both the inputs. I am going to draw a 2-input NAND gate this is V_{dd} and then the other one is ground and here we have the PMOS transistors and then the NMOS transistors.

The sizes which we had used was 2 and 2 and then here also 2 and 2 and if I consider this as A and A and then this as B and B. We know that the input capacitance of A and B remains the same input capacitance for A is 4, the input capacitance of B is also 4. Now, let us take another variant of this 2-input NAND gate, I am going to write it as 2-input NAND.

I am going to have the same design in the form of the transistor level, it is nothing but 2-input NAND gate, it will have the same design. I am going to say that on the PMOS side I am not going to change the sizes. But on the NMOS side I am going to change the size and I am going to say that this is 4/3. But I want this falling resistance to be the same as that of the falling resistance of the 2:1 inverter, 4/3 is likely to give me a resistance of R/(4/3), which will be nothing but 3/4 R. If I choose a width of 4 here this one will give me a switching resistance of R/4. The total of this switching resistance is the sum of the switching resistance and this switching resistance will give me the switching resistance of R. Which is what we want to compare the logical efforts or the parasitic or when we want to estimate a logical effort or the parasitic with respect to 2:1 inverter. What we have done here is I have to change the widths of the NMOS transistors for the 2-input NAND gate. Up till now we had seen the same sizes, now we have changed it into a different width. But giving the same falling resistance, here also the falling resistance was R.

But the widths of the transistors are not symmetric anymore that means the input capacitance seen by the input A and the input B are asymmetric, that means they are different. The input capacitance seen by the A input will be nothing but 2 + (4/3), the input capacitance seen by the B input which will be nothing but 2 + 4 = 6C, hope this is clear, moving on.

(Refer Slide Time: 24:40)



Now, why do we do this having an asymmetric gate, input capacitance seen by inputs are different. That is what it means, on one side we have a larger input capacitance for this particular transistor and then a smaller capacitance for this particular transistor.

Why it is used is? If we have A and B and one is a data signal. This being A data signal and another one is a control signal. Let us say that the B is a control signal. The control signal is of low frequency in the sense it is operated not every time.

It is operated at a low frequency, once in a while it will come into the picture, it will be enabled. That is when the B is going to get operated, the B is going to get a signal of 1 rarely or at a low frequency. In that sense, now I will have a capacitance associated with this 4C, because it occurs very rarely this 4C, when the output to fall down from 1 to 0, even this particular capacitance will discharged.

This capacitance being more, so 4c is actually derived from the width of 4 here and because I am increased this width this capacitance also increases and then this particular capacitance has to discharge. This particular very rarely occurring signal, whenever it happens this 4C capacitance has to discharged.

This particular data signal is a high frequency signal that will more occurs more frequently, this particular signal incorporates the transistor of the size 4/3, that means that the output capacitance is nothing but it incorporates 4/3 capacitance.

Of course, this 4/3 will get added up with this particular 2 capacitance. But what I am saying is if I can change this particular width here, it is going to accommodate to the output node a lower capacitance. My output node capacitance is likely to decrease which will lessen the parasitic. Also the logical effort of this particular data signal which is the A signal connected to this particular transistor that is also going to get reduced.

This 4C capacitance also likely to take more amount of time to discharge, but because it is a rare signal it is a low frequency signal that is perfectly fine when compared to a mostly occurred signal. whenever we do this size increase in the size of the transistor.

This particular transistor input is also made available, it is also connected to the early arriving signal. The reason is if suppose I have two signals A and B, if the B arrives early then it becomes the 4C capacitance here can easily discharge, if B is 1 then the 4C can start discharging, it does not have to wait for the A signal to arrive.

In that sense, if among these signals if we have calibrated or characterized in such a way that we know that among these two signals B is arriving always early, then I can easily increase the width of this transistor during my designing stage. So, that even if I have a very high capacitance it will get discharge early, because B is arriving early.

At the same time if I increase the size of the B transistor here, I can compensate with having a lower size here, that we will get an equal falling resistance of R and a lower size on this particular transistor will not only help in decreasing the parasitic, but also the logical effort for the transistor A.

Let me try to find out what should be the logical effort of the transistor A which is nothing but,

$$g_A = \frac{2 + 4/3}{3} = \frac{10}{9} < \frac{4}{3}$$

Because I am having a 2 is to 1 benchmark inverter which is seeing the same output current, because of the same falling resistance or the same rising resistance.

10/9 is actually less than 1.33. What about the parasitic? The parasitic here, because 4/3 is a less size, a less capacitance get incorporated into the output node, we now have,

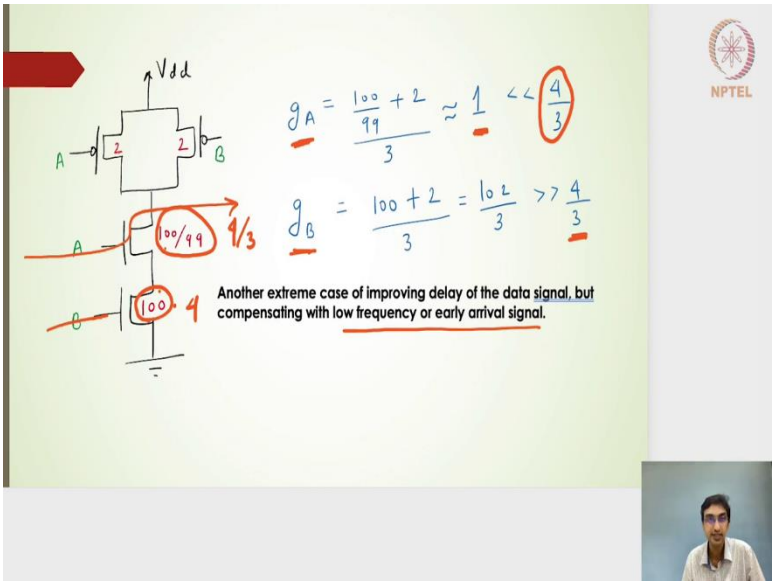
$$P = \frac{2 + 2 + 4/3}{3} = \frac{16}{9} < 2$$

Both my parasitic and then the logical effort according to the linear delay model has decreased. Where we have compensated into the other transistor which is B and naturally the other logical effort of the B transistor is going to increase.

$$g_B = \frac{4 + 2}{3} = 2$$

Then 2 is greater than with the 4/3 for the 2-input NAND gate. But this is because, this is well accepted because this B transistor is either it is a rarely occurring signal, so it is a low frequency signal. Only at that rare event we will use this particular logical effort to find out that particular delay. Although in that rare event we will have a larger delay because of this B transistor path. But overall, most frequently the A signal is our data signal. This particular path is going to have a less delay because of the less logical effort and then less parasitic, hope this is clear.

(Refer Slide Time: 29:58)



What if I make it to more? Let us instead of 4/3 and 4 here, if I make it 100/99 and 100 here. I have a huge width here on the B transistor and A is nothing but a 100/99. Naturally the logical effort of A will be nothing but,

$$g_A = \frac{\frac{100}{99} + 2}{3} = 1 << \frac{4}{3}$$

Logical effort of B is

$$g_A = \frac{100 + 2}{3} = \frac{102}{3} \gg \frac{4}{3}$$

If it is really a low frequency, we can try accommodating this, but remember that 100 a size of 100 is going to increase your area drastically when compared to a size of 2:2.

If I consider a 2-input NAND gate with the NMOS size of 2 and the PMOS size of 2 or NMOS size of 2 and then PMOS size of 2 we used to have the size of 2 is very very less compared to that of the 100. If we really want this particular data path to have a very high performance, then only we increase this particular size. Otherwise, we have it 2:2 or we can try to accommodating 4/3 here and 4 here.

This is another extreme case of improving the delay of the data signal, but compensating with only the low frequency or early arrival signal, which is connected to this particular B transistor. Hope this is clear to everyone. I am going to stop my lecture here.