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

Lecture - 14
Equivalent of transistors in series

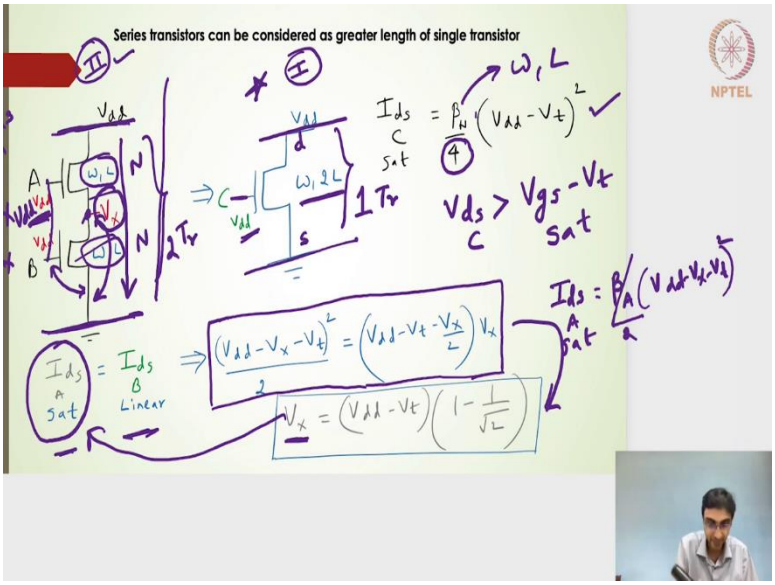
(Refer Slide Time: 00:16)

Welcome to this particular lecture. This lecture we will talk about or we will discuss more on the CMOS simplified or the simplification of the CMOS transistors in terms of the width and the length scaling. Especially if you have 2 or 3 CMOS transistors, whether it is an NMOS transistor or a PMOS transistor. How do we get an equivalent simplified circuit and then estimate it is current or whatever the output node voltage?

We will look into one such example in this particular lecture and then the other part of the lecture we will talk about the pass transistors. Which is also kind of a called as a different family, but it is the fundamentals or the concepts is actually derived from the CMOS transistors.

We will look into the pass transistors, as well as finally we will have a look into the transmission gate, which is nothing but made up of the NMOS and PMOS transistors. Till now what we had seen is the long channel current model, the short channel current model for the NMOS and then the PMOS transistors and also in the last lecture we had realized the inverter DC characteristics and then the transfer characteristics.

(Refer Slide Time: 01:40)



Let us say that we have one transistor here, this is what I am pointing to where one transistor with the input is V_{dd} and it is an NMOS transistor there is no bubbled gate here, that is why it is an NMOS transistor and its a standalone transistor. What I mean by a standalone is, it is not connected to any other part of the PMOS transistors or any other part of the circuit. It is a single NMOS transistor where one side it is connected to the V_{dd} and another side it is connected to the ground.

What it implies is, on one side which is connected to the V_{dd}, it should be the drain terminal, the other one which is connected to the ground, it should be the source terminal and this is the input it is supplied with the V_{dd} value. We have one more set of configurations, where we have two NMOS transistors, even this is NMOS and this is also NMOS transistors and they are in series.

Again, just to reiterate, it is not PMOS and NMOS it is both of them are NMOS. Two of the NMOS transistors in series and one is connected to the V_{dd}, the other terminal is connected to the ground and in between terminals are tied together. The input to this particular first transistor is an A transistor and the input to the second transistor is a B transistor and both of them are supplied with the V_{dd} potential.

Basically what it means is, this particular configuration the inputs are tied together and then they have been given the input of V_{dd}. Now, what we have is one configuration, here the first configuration and then the second configuration. The first configuration the input

is V_{dd} and we have a transistor configuration although the number of transistor configuration is only one transistor and then this is a one rail this is another rail.

We have a V_{dd} rail and then a ground rail and then here it is only one transistor. I am going to write it as 1 Tr (1 transistor) the input is V_{dd}. We have now the input supply as V_{dd} the two rails has one transistors in the configuration the rails are V_{dd} and 0 rails.

Similarly, here in the second configuration we have two rails V_{dd} and then ground, the input is V_{dd} and the rails are V_{dd} and ground and then here we have nothing but two transistors (2 Tr). Two configurations are there.

Let us see what is the overall current that is flowing from the configuration 1 and configuration number 2. I am calling this particular configuration, configuration number 1 as the C transistor. That is why I have annotated or the input is this C, the second configuration the inputs I have annotated it as A and B, the reason is there are two transistors, one transistor we will label it as A transistor, the another one transistor is we will label it as B transistor. This one the first configuration we will call it as the C transistor.

The first one out of the two configurations, the first one is kind of very easy to realize or understand or write down the current equations. Let me begin with writing the current equation for the first transistor a first configuration.

I am writing I_{ds} for the C transistor and I am writing it as saturation region. Let us see whether this transistor will be in the saturation region. If I look into the terminals drain and source and then the input C or rather the input is V_{dd}.

The V_{gs} here is V_{dd}. V_{ds} value, V_{ds} for the transistor C is actually greater than V_{gs} (which is also V_{dd}) $- V_t$, let us say we can consider it to be 0.3 volts. In this case, if this satisfies the transistor will be in saturation.

That C transistor is in saturation is given as,

$$I_{ds_{C_{sat}}} = \frac{\beta_N}{4}(V_{dd} - V_t)^2$$

Now, let us begin with the second configuration. The second configuration we have the transistor A the transistor B and then we have supply or the rails are V_{dd} and then ground.

In between there will be a current supply here, if there is a current supply, if there is a current that is flowing from the V_{dd} to the ground, then there will be a potential called as V_X.

This $V_{dd} - V_X$ that will be the drain to source, which will enable the current to be driving in this particular transistor A. Similarly, this V_X to ground will enable the current to be driven in this particular drain to source of the transistor B.

Now, let us see in this particular transistor A. If I write it here transistor A on one side it is V_{dd} and on the other side it is V_X potential. The current here is I am going to write it as I_{ds} of A and what should be the transistor operating region, whether it should be in the linear region or whether it should be the saturation region.

The input applied to the transistor A is nothing but V_{dd}. If that is the case, the V_{gs} value for the transistor A is nothing but $V_{dd} - V_X$, the V_{ds} value is also $V_{dd} - V_X$. The transistor here in this case will have or will be operating in again saturation region. The transistor A I am talking about will be operating in the saturation region, because $V_{ds} > V_{gs} - V_t$.

That is what I have written here the I_{ds} current for the transistor A is in the saturation region. Transistor B if I look into closely the input is V_{dd}, V_{gs} is nothing but V_{dd}. V_X is there, which we do not know what is the value of the V_X. V_{ds} for the transistor B here is $V_X - 0$.

Is it possible that the V_X value can be greater than $V_{gs} - V_t$, that is the question. What I meant was a transistor A should supply some current, the same current should go into the transistor B, because both of them are in series. That is why I have stated that transistor A current should be equal to transistor B current.

Although I have written it as linear, let us see why it should be in the linear operating region and not in the saturation operating region. For this transistor B, if it has to be in the saturation operating region, V_X value should be greater than $V_{dd} - V_T$. The moment V_X becomes $V_{dd} - V_T$ here, $V_{dd} - V_T$ what it implies is this is $V_{dd} - V_t$, transistor A input is V_{dd}.

V_{gs} for the transistor A turns out to be less than V_t. Let us pick some particular values, V_{dd} is 1 volts, V_t is 0.3 volts, if V_X is actually more than $V_{dd} - V_t$. That means, if it is not

0.7 or 0.6, it is actually more than 0.7, it is 0.8 volts input to the transistor A is 1 volt. V_{gs} for the transistor A turns out to be 0.2 volt, transistor A will be in cut off. If I want transistor B in the saturation mode, then transistor A turns out to be in the cut off.

There will not be any supply of the current right in our initial assumption, what we had said was V_X will reach to a value, such that it will drive the current from drain to source for the transistor A and then it will drive the current for the transistor B also there. There will be a non-zero current that will be flowing, that will determine the potential of V_X.

If I want some nonzero current, this V_X value cannot go beyond 0.7 volts. That means, it cannot go beyond $V_{dd} - V_t$, it has to be less than $V_{dd} - V_t$. If that is the case transistor B will always be in the linear region. The reason is, if V_X is always less than $V_{dd} - V_t$ then V_{gs} for the transistor B is $V_{dd} - 0$ and V_{ds} here will be for the transistor B will be less than $V_{dd} - V_t$.

Thus V_{ds} value will always be less than $V_{gs} - V_t$ and thereby the transistor B will have a operating region is linear region. If I now equate these two current equations V_{dd} and whatever the β of this particular transistor multiplied by I mean if it is the saturation region.

I am going to write, let me write it here. I_{ds} of the transistor A has to be in saturation.

$$I_{ds_{A_{sat}}} = \frac{\beta_A}{2}(V_{dd} - V_X - V_t)^2$$

Then Ids of B will be,

$$I_{ds_{B_{sat}}} = \frac{\beta_B}{2}[2(V_{dd} - V_t)V_X - V_X)^2$$

That will be the linear region transistor B current. I will have actually β_A and β_B, the dimensions of this transistors which is given as W L, W comma L are same. If the dimensions of the two transistors are same, assuming that the t_{oxide} is also the same here, the mobility of both the NMOS transistor for a given particular technology node is also the same.

Now I have the β for the transistor A and β for the transistor B to be actually same. Then I can actually equating the currents for the A transistor and B transistors. I can actually

neglect the β and what we have is nothing but this particular expression. We will have these particular expressions,

$$\frac{(V_{dd} - V_X - V_t)^2}{2} = (V_{dd} - V_t - \frac{V_X}{2})V_X$$

When we equate the current for the transistors with the same dimensions of W and the length, the channel length.

From this I should be able to find out, what is the V_X node voltage value V_X. V_X turns out to be,

$$V_X = (V_{dd} - V_t)(1 - \frac{1}{\sqrt{2}})$$

$$V_X = (1 - 0.3)\left(1 - \frac{1}{\sqrt{2}}\right) = 0.2099V$$

V_X turns out to be 0.209 or 0.21 volts, it satisfies our conditions of saturation and a linear region. For the second configuration we have got the V_X value. Let me once again go back to this particular configuration number 1 here the scaling or the dimensions is W comma 2 channel length. The channel length of this particular transistor is twice the channel length.

Let us say that, this particular configuration number 2 is designed on a particular chip and then the configuration number 1 is designed on a particular chip. Two separate chips. Generally, when we are manufacturing on a particular chip or a particular wafer, where we extract multiple chip designs, we tend to have the same channel length.

But because I have written two different channel lengths, here channel length of L here for the configuration 2 and channel length of 2 L here in the configuration number 1. What I am just to make it simplified those two chips or the wafers are completely fabricated separately.

We have this two-channel length and this is the channel length of one channel length in the configuration 2, in the configuration 1 it is a 2-channel length with the W. The current of this particular transistor will be, $\frac{\beta_c}{2}(V_{dd} - V_t)^2$.

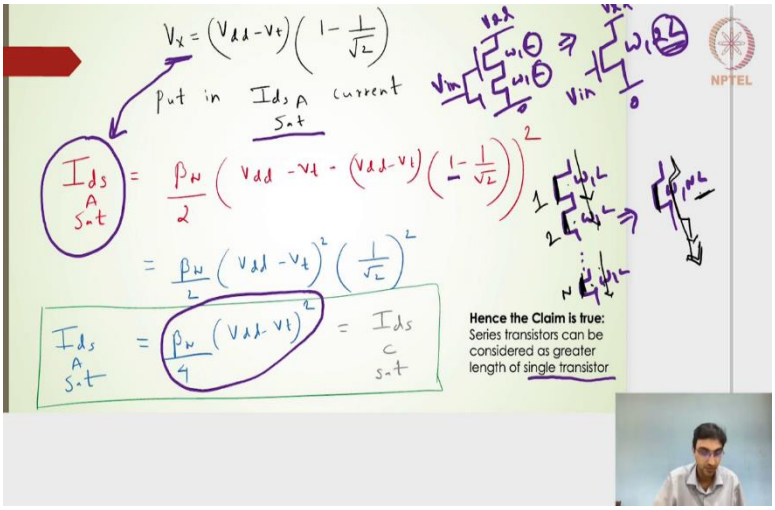
But if I want to write this equation in terms of the β of the transistor A and B, where the β were same, it was for the channel length of L and that is why the βn here, which represents the transistor A or transistor B is a β value. That I have written here because it has a 2 channel length here. That 2 additional 2 will come here, in the denominator and that is why it becomes $\frac{\beta_N}{4}(V_{dd} - V_t)^2$.

This β_N is for the transistor with the dimensions of W, L. Since the first configuration has the channel length of 2L. I am writing this as β_N which is of for the W, L and then divided by 2, because you know if you remember the β is a function of W/L. L is in the denominator side, if the channel length increases, the β actually decreases with respect to W/L it will be β/2 and then additional 2 is coming because of this current equations. Now, if I coming back to configuration number 2, now if I want to find out what is the actual current of this particular configuration 2 when transistor A and transistor B are in series. If I can pick one of the current and put this V_X value, I should be able to find this particular current.

Now I have V_X value, I am going to input this V_X value into the transistor A saturation current equation, saturation current generally is much more simple than that of the linear current and that is why I am going to put this V_X node value into the saturation current equation.

Hope you have understood this.

(Refer Slide Time: 18:01)



Going to the next slide, if I put this V_X value which we have got from the previous derivation. Put this in the current equation for the transistor A the saturation current. The saturation current for the A will be nothing but,

$$I_{ds_{A_{sat}}} = \frac{\beta_N}{2}(V_{dd} - V_t - (V_{dd} - V_t)(1 - 1/\sqrt{2}))^2$$

This 1 is there which will get cancelled with this one and then I am left with,

$$I_{ds_{A_{sat}}} = \frac{\beta_N}{2}(V_{dd} - V_t)^2(1/\sqrt{2}))^2$$

$$I_{ds_{A_{sat}}} = \frac{\beta_N}{4}(V_{dd} - V_t)^2 = I_{ds_{c_{sat}}}$$

If you notice this particular expression, which we have derived for the current for the transistor A by putting this particular V_X value into this particular transistor a current turns out to be same as nothing but the transistor C current.

What we are saying is the series transistors, if I have 2 series transistors, 2 transistors of the same dimensions W, L and then the next transistor is also W, L. The series transistors can be considered as a greater length of the single transistors. Its equivalent, what we did was, I had 2 transistors, NMOS transistors W, L at the same input here, this could be same as nothing but having one single transistors off twice the length with the same input.

If input is V_{dd} here, the same input should go here V_{in} and of course the same rail should be there V_{dd} and ground, then we can simplify these 2 series transistors into an equivalent single transistor of W, 2L. Where this channel length is considered as the length of the single transistor.

This is nothing but equivalent 2L is coming from this L plus this L. In fact, what we can actually do is, if I have two transistors if one transistor has a length of L, the another transistor has a length of 2L, the equivalent transistor can be considered as W, 3 L.

In fact, the one more generalization which we can do is, if I have N such transistors with W, L; W, L; W, L, I can simplify this into a single transistors of W, N L. This is one such example showing a simplified 2 series transistors or the 2 transistors which are stacked,

can be simplified to get the current expression by replacing it or substituting with an equivalent single transistor, where the length has been increased.

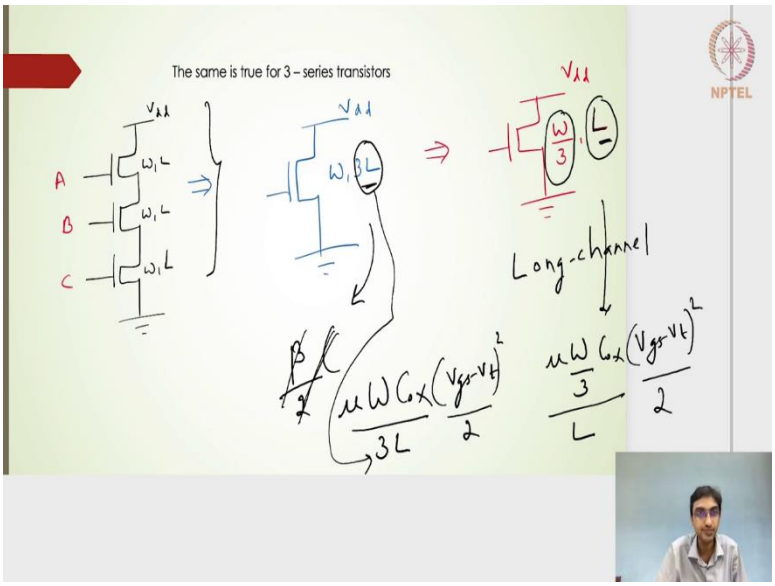
An intuitively it makes sense because if I am talking about this particular single transistor channel here and then I am talking about this particular channel here and if I am talking about this particular channel here, if you remember the current for the equivalent.

I am talking about the channel of this. The current has to pass through this particular channel here, the current has to pass for channel 1, channel 2, channel N. This is a current that has to see the same resistance, which is N times here for an N transistors configuration, the N transistors which are actually if I say 1, 2 and N transistors, for an N transistors which are stacked together or which are in series the current has to actually see a resistance of N times the single transistor channel resistance.

Similarly, the equivalent one should have an effect, this current should also see, should also undergo a channel resistance N time that of the single transistor channel resistance. That is why if I make this N times the channel length, the channel resistance will be N times. Intuitively it makes sense.

For the 2 transistors also, we have got this particular example, which validates the two series transistors equivalent transistor will have the twice the or rather the channel length of the equivalent transistor is nothing but the summation of the 2 transistors which are in series.

(Refer Slide Time: 23:31)



The same is true for the 3 transistors also. Here also I have one configuration where we have 3 transistors W, L of the same width channel length is also same. The same dimensions W, L; W, L; and W, L and what we notice is that we can actually replace or substitute for simplification purpose, we can actually have a single transistor of W, 3L.

In fact, we can also do for a long channel model. This is applicable only for the long channel current model. We can also say that, it is W/3, the width gets divided by 3 comma the same channel length, the reason is very simple, if I write this particular equation. The current equation, which is nothing but the saturation equation, β by 2, let me try to write it in terms of not in beta. Instead of that, I will write $\mu WC_{ox}/3L$ because this is the 3L here and then we have this $(V_{gs} - V_t)^2$.

The current for this the current for this, the long channel current for this particular transistor also is $\mu WC_{ox}/3$, of course there is a 2 here. $\mu WC_{ox}/L$ and then I will write $\frac{(V_{gs}-V_t)^2}{2}$.

If I look into both these expressions, the $\frac{w}{3}$, L and W, 3L both of them turns out to be the same. I can also have a dimension instead of channel length being extended to 3L. I can also have it $\frac{w}{3}$ as the width, just for our analysis purpose we can as well take W, 3L or we can take $\frac{w}{3}$, L. If I have this 3 series transistors or stacked transistors right, hope you have understood this.