Digital Integrated Circuits
Dr. Amitava Dasgupta
Department of Electrical Engineering
Indian Institute of Technology, Madras
Lecture No # 28

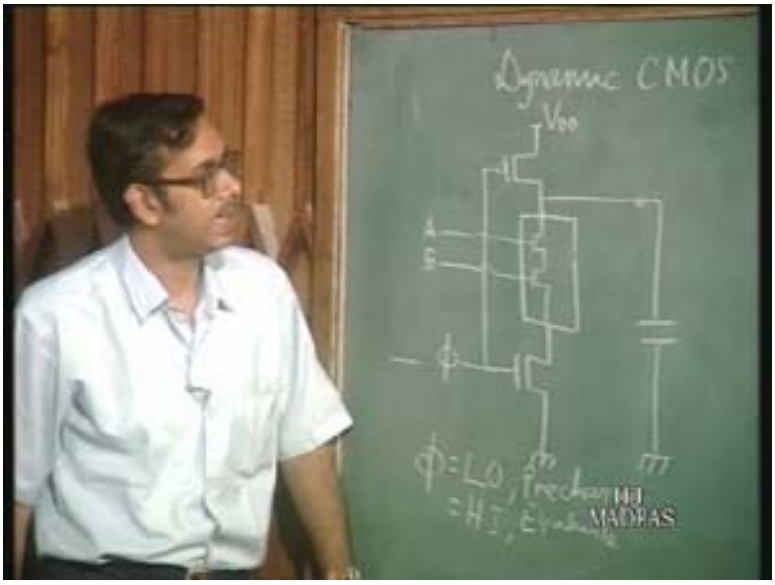
Dynamic CMOS; Transmission gate;
Realisation of MUX, decoder, D-F/F

We shall continue our discussion on MOS logic circuits. The circuits we were discussing last time was the dynamic CMOS logic circuits or clocked CMOS logic circuits. So we had seen that if you have a circuit like this, that is you have a PMOS then an n block and then an NMOS transistor and you have a clock which is connected to the circuit and this is the n block which consist only of NMOS transistors and the inputs are fed to the n block. For example if you have NAND gate, we had taken that example last class. That is you have 2 NMOS transistors here then when the clock is low what happens is this NMOS is off and the PMOS is on and you have a load capacitance say which may be the input capacitance of the next stage gates.

What happens is the output voltage is charged to V_{DD} so this is called the pre charge phase. So when phi is 0, phi is low say you have the pre charge phase you are pre charging the output and then when phi goes high what happens is the PMOS goes off. So there is no connection to V_{DD}, the NMOS turns on so there is a possible path to ground provided that these transistors, the n block has a block provides a connection to ground.

So in this case if you have two inputs A and B, only when A and B is high then only there is a path to ground. So when A and B is high then only the output is 0. This is a NAND gate, so the capacitance gets discharged. So when phi is high this is the evaluate phase. That is when phi is high the output, basically you generate the logic function in the evaluate phase that is when phi is high. So this is a clocked CMOS circuit. The obvious advantage is that here whatever be the complexity of the gate, if you have say n number of inputs, in a normal CMOS you will have to twice n number of transistors whereas here it is going to be just $n + 2$.

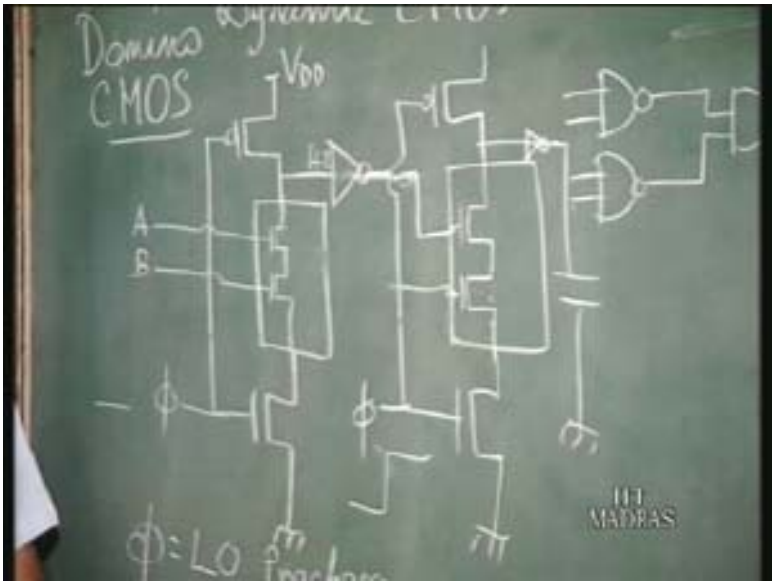
(Refer Slide Time: 04:30)



So basically here you have the lower n block of a CMOS. In a CMOS you have the lower n block and the upper p block of transistors, so here you have just the n block of the CMOS. So this is dynamic CMOS but this circuit has certain disadvantages suppose for example if this particular circuit is driving a similar dynamic CMOS gate, for example you have another say n block and you have the similar circuit, another clock here, the same clock is fed here and this is going as input to this n block where you have NMOS transistors.

So what is going to happen is when the clock is low that is the pre charge phase, this input capacitance of this transistor here in the second stage that is going to be charged to V_{DD}. Suppose this clock goes from low to high so for this gate you are going to the evaluate phase from the pre charge phase. So at this instant when the clock goes from low to high provided that again you have a NAND gate and this inputs are all coming from previous stages, all the inputs to these gates are going to be high.

(Refer Slide Time: 10:35)



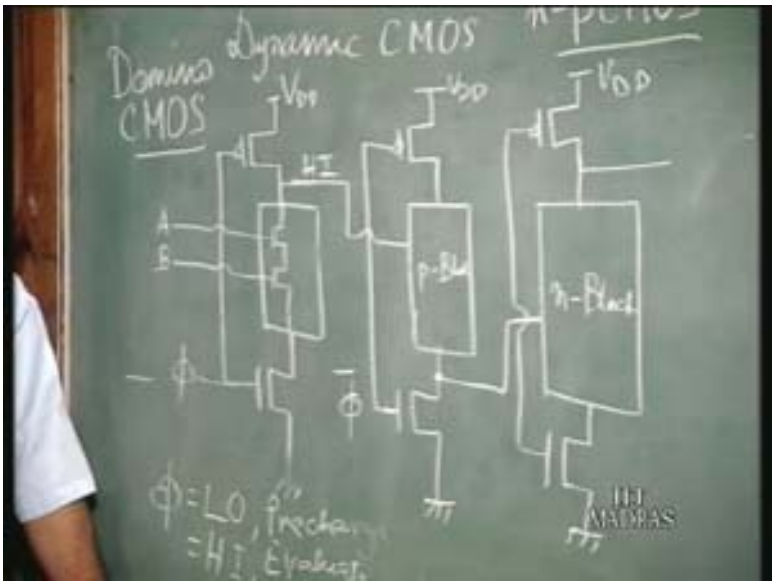
All the clocks go high simultaneously. So before this input has time to react to the changes in the previous gate, you know first this output must change and then effectively the other output must change. Basically if you have something like this, you have a NAND gate say two NAND gates and then it is like this. So depending on the inputs here, these outputs will change. Once this output changes then depending on this which is the input here, the next output is going to change. Basically what happens here is these inputs are high at the time when the clock goes high. So before it has time to react what is going to happen is again in the pre charge phase, you have to charge this capacitance. So there is a path to ground immediately because these transistors are all on.

So this capacitance is going to get discharged because this inputs are all charged to high, all this transistors are on at the beginning of the evaluate phase. So that is going to create problems because once this charge is discharged there is no way in the evaluate phase, the output can go high. So the way to circumvent this problem is we have to ensure that at this instant when the clock goes high that is the start of the evaluate phase all this transistors in the n block they must be off. How do you do that? We just put an inverter here. An inverter is introduced here at the output of this circuit. What happens is basically now when in the pre charge phase, the input capacitance of the inverter is charged to high. The output is always low and this output is going to the inputs of the next stages.

In the pre charge phase all the inputs of the n block are low. So all this transistors are basically off, all the NMOS transistors in the n block are off. So this is true for all the gates here. now what happens is suppose we have stages like this, suppose A and B goes high then what happens is the output which was high here in the pre charge phase, this capacitance gets discharged through this path. So this high goes low then the output of this inverter goes high. So then only this transistor turns on and suppose this input is coming from a similar gate then only it has a chance to turn on and once it turns on then only this capacitance that the input of this inverter can get discharged.
So the input of the inverter was charge to $V_{DD.}$ during the pre-charge phase then only it gets chance to discharge. So you see what happens is first this output of the first stage will change and depending on that the next state will change like that. So this particular circuit is called the domino CMOS circuit because the changes happen one after the other. First state changes then depending on that the next state stage will change and so on and so forth. So this is the circuit of a domino CMOS. Now one can really avoid this inverter by using what is called an n-p CMOS. So here what you do is the next block is basically made up of PMOS transistors. The circuit is going to be something like this. Again what you have is a PMOS transistor here, this is actually a p block not an n block that is consisting of PMOS transistors and this you have the NMOS transistor.

So here instead of phi, phi bar is applied, the compliment of phi is applied and so this output here can be fed directly as input here and then again you have the n block. So the next stages are all n block so this is VDD, so the output comes from here, goes to n block again you have similar; this is PMOS and NMOS so this output is taken from here. So what happens is if you just look at it now when phi is low, it is again the pre charge phase.

(Refer Slide Time: 16:40)



So what is going to happen? All the outputs capacitances are going to be charged. So the output is just the pre charge value, it is not the exact output of the logic circuit. So in the pre charge phase when this goes low, here all the outputs are high and this is going to the p block. All the PMOS transistor inputs are high, so all the PMOS transistors are off. Now here what happens is this is you have applied phi bar, so when phi is low, phi bar is high. This NMOS transistor is on. The output of the p block is low. The inputs of the NMOS n block are all low.

All the NMOS transistors in the n block they are off. Basically it is like this that to ensure that this works properly, at the beginning of the evaluate phase, all the transistors in the p block or the n block they must be off and then depending on the inputs they can turn on. If they turn on then only the output changes otherwise it retains the pre charge value that is whatever was the output during the pre-charge phase. What happens now is suppose if you have this configuration here, in the evaluate phase the phi goes high that means the NMOS transistor here turns on.

Depending on this input here, the output can be discharged. Discharged means the output here, so it goes to the input of a PMOS transistor. So that PMOS transistor input can go low which means that PMOS transistor can turn on and so if the PMOS transistor turns on, you say there are number of PMOS transistors in this p block and suppose this p block as a whole turns on that is it is conducting, so what can happen is this output was basically low.

If this is conducting, so in this phase this NMOS is off and the PMOS is on because this is phi bar. If this is conducting, the output here this is off (Refer Slide Time: 14:50). So this can charge to $V_{DD.}$ so this output of this p block here can go up to $V_{DD.}$ which means that this can turn on the next n block transistors. So again it is like a wave, so

first these transistors have to turn on and depending on this, this can turn on the next block transistors and then if it turns on the next block then the next block can turn on and the output can change. So this type of logic gates are clocked logic gates and the advantage here is of course that you save a lot in terms of transistors and equivalently in terms of area. So these are quite popular nowadays.

One disadvantage you can say is that one point to be noted is that in this case there is a limit on the lower frequency of operation that is you can operate the circuit above a certain frequency. This clocked frequency cannot be made too low. The reason is like this that it basically works on the principle that during the pre-charge phase, you are charging the output capacitance of this gate which is the input capacitance for the next stage and then if this block is not on, the charge is retained and so the output voltage retains its previous value but all this capacitances are not ideal capacitances there are leakage paths and so the output charge is going to decay with time. So this pre charge evaluate phase must be quick enough so that before the charge decays you have the next pre charge and evaluate. This evaluate phase or the pre charge cannot be too long otherwise the output voltage is going to change.

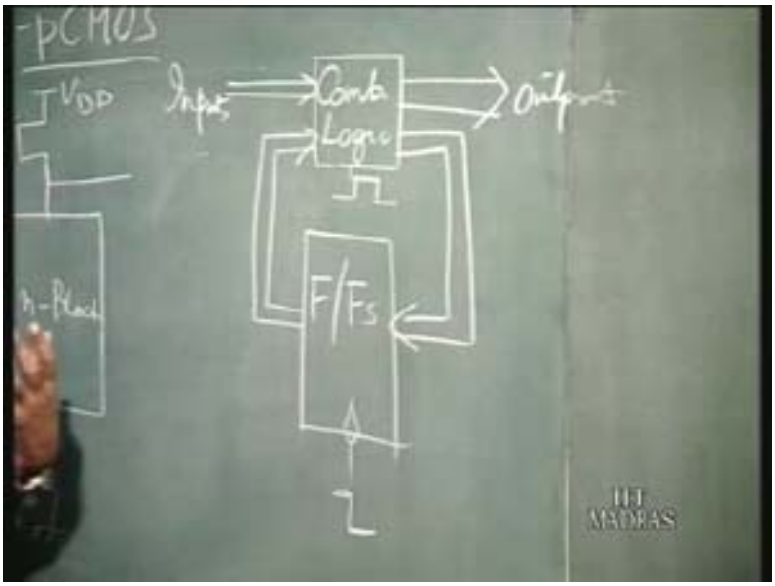
For example this p block input what happens is when you charge this, this input goes high in the pre charge phase and then suppose this NAND gate one of the inputs is zero. So this is off, this p block input which was charged to high is going to remain high because there is no path to ground. That is the expected output of this logic circuit but because of some leakage paths this output can decay with time. If the evaluate phase is too long, slowly the output voltage is going to fall. The period has to be short.

So all this circuits, they have a lower limit on the frequency of operation. That is you have to operate it above a certain frequency which is not really a big disadvantage because nowadays the clock frequencies of operation are quite high and so this is not really a disadvantage. I just like to say that this type of circuit, although you have a clock and it may seem that you actually require an additional resource here to get this circuit working but it is not really so.

See if you look at any system digital system which consist of combinational sequential logic circuits for example if you take the finite state machine model which can be used for any digital system say what you generally have is you have a set of inputs then you have a set of outputs. This is the combinational logic and then what you have is you have a set of flip flops. So from here you generate the states of the system and you feed it to the flip flops and from here this goes in as the input to the combinational logic. This can be treated as a model of a digital system. So basically here in the flip flops, you require a clock and this clock can be a negative edge trigger flip flop, so the changes take place in the negative edge of the clock pulse, when this of course goes from high to low.

Now suppose here you have a digital circuit, a dynamic logic circuit where the changes take place, the evaluate phase is when the clock goes high. The similar circuit as we have already seen. So what happens is when the clock is low, it is in the pre charge phase. So the output comes here but it is not going to affect the flip flops because the clock is low. Now when the clock goes high, you have the evaluate phase that during this time, the output of this combinational logic block is going to change depending on the inputs to that combinational block. So the output is ready when the clock is high and then what happens is then the clock goes from high to low then depending on what is the inputs, you see the output of the combinational block is actually fed as the inputs to the flip flops. So when the clock goes low, depending on the inputs to the flip flops the output changes but then the clock is low.

(Refer Slide Time: 22:30)



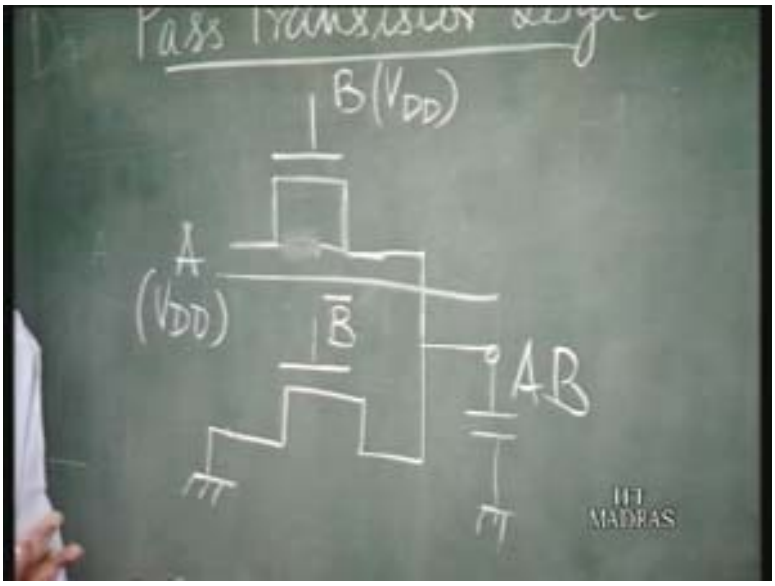
So when the output changes here but the clock is low. So again this is the pre charge phase but the output of the flip flops is retained till the next negative edge of the clock pulse. So this input will be here ready and again you have the inputs to the system. So when the clock goes high, again you have the inputs ready so depending on that again the output of the combinational logic is going to change and then it is again going to affect the flip flops. So you see that it is a perfectly viable system you use the same clock for the combinational logic part as well as the flip flops. So basically there is no extra overhead actually for having this clock here.

So the same clock can be used for the combinational logic. What you gain here is if you use the normal CMOS circuit, you would require twice the number of transistors, here you require almost half the number of transistors and all the advantages of CMOS are still there. So you don't lose out anything. This is actually a very popular form of circuit which draws the advantages of the CMOS logic. At the same time the disadvantage of CMOS of having too many transistor is reduced.

So the next form of circuits which we shall take up for discussion regarding MOS logic is the circuits which come under this category of pass transistor logic. What is the pass transistor? A pass transistor is nothing but say an NMOS transistor. Now why is it called a pass transistor? It is called the pass transistor because when the gate voltage of this NMOS is high then it behaves as an on switch. You can consider this as a switch basically in a pass transistor logic we consider this as a switch. So when the gate voltage is high, this transistor is on, it's just like a switch. So whatever is the input is passed on to the output and when this gate voltage is low, this behaves as the off switch so there is no connection between input and the output. That is the basic property of this. Now this can be used to realize logic gates.

For example if I say that this input is A and here this is B, what is the logic at the output? If B is high then the output is A. But that does not give us all the conditions so basically what you have to do is we have another transistor here, this input is the compliment of B. So let's draw it properly, so this is B bar and we make this go to ground.

(Refer Slide Time: 28:27)



Now what you have? When B is high, output is A that means if A is also high, output is one. When both are high then only the output is one, if B is high A is 0, output is 0. When B is zero then this is off, this is on (Refer Slide Time: 25:46). So output is 0, so only when A and B are both high then only the output is high so it is a AND gate. You have realized an AND gate with the pass transistors.

The advantage you can see very well is that you have just use two transistors and we can also say that the power dissipation effectively does not consume any power because it's not connected to power supply. It is just using the power inputs. So it is very low power dissipation circuit as well as advantage is you have very high density in the sense that you can realize logic functions using very few transistors. So both these advantages are what you require for large scale integration. This concept of logic gates is ideally suited for large scale integration.

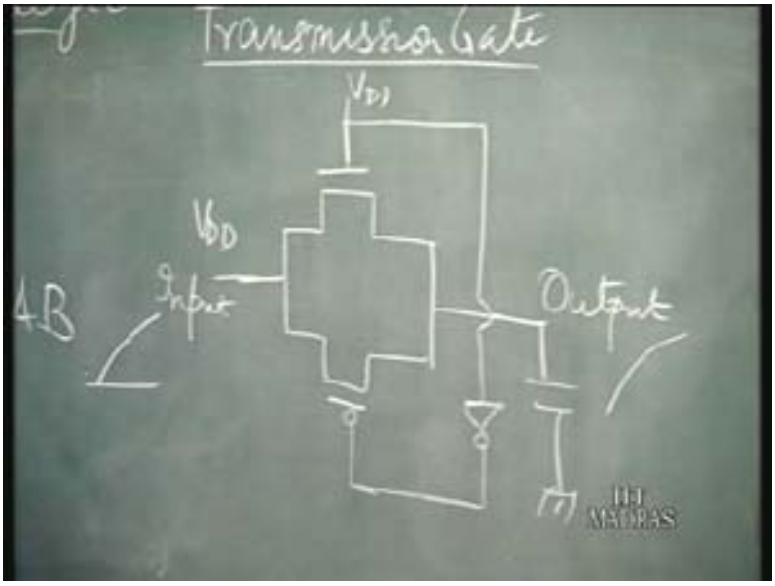
So this is one example we have given but always we must be careful when we use this for logic circuits, we have seen this. We cannot just say this is A and B because when B is zero if we just use this block here, when B is zero the output would be undefined. So you must always take into account that all the possible input combinations are accounted for. So that is one constraint we have to think of. The other point is suppose this B is high here, so that is this high means $V_{DD.}$ and suppose A is also high $V_{DD.}$.

Now what is going to be the output? Initially say output was zero so when this turns on, the output again we think of that as some capacitance here. So this output has to be charged to $V_{DD.}$ so the output rises. Basically you can consider this as a resistance here, this transistor we can consider it as a resistance. When this transistor turns on it is a low resistance and when this transistor is off, it is a very high resistance.

So when this is a low resistance, this capacitance is going to charge with an R_c time constant slowly it approaches V_{DD} but when it reaches V_{DD} minus V_T what happens is so this is basically the drain and this is the source because this is a higher potential. So when this source voltage reaches V_{DD} minus V_T then this transistor cuts of because it cannot rise any further because it cannot supply any more current once the gate to source voltage falls below V_T. So there is a drop of V_T for this pass transistor. So the output voltage can only rise till V_{DD} minus V_T or V_G minus V_T whatever is the gate voltage, so this is a drop. So you have better switches to realize this same logic, what is called the transmission gates.

A transmission gate consist of an NMOS and a PMOS. This is an NMOS, you have an inverter here, the two gate voltages are compliments of each other and suppose this is the input, the other one is the output. This is the control input here, now suppose you have a high here V_{DD} as we had in the previous gates or the gate of the NMOS. The voltage at the gate of the PMOS is going to be 0 or ground say because this is the output of an inverter coming here. So inverter input is V_{DD}, the output is going to be ground.

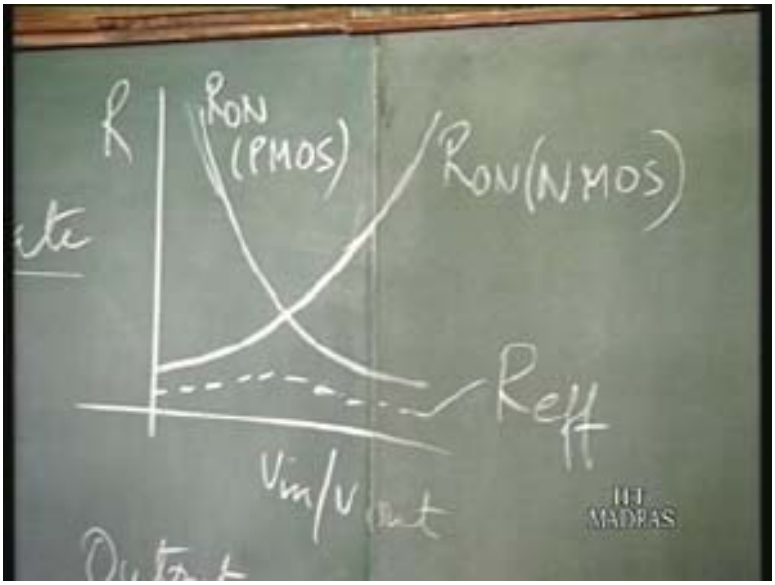
(Refer Slide Time: 32:32)



What happens is suppose you have an input here high then what happens? This is V_{DD}. What was happening was as this output voltage was charging up, the gate to source voltage of the NMOS was reducing. So that the effective resistance of the NMOS transistor is increasing. It requires more and more time to charge this output capacitance and finally when the output voltage reached V_{DD} minus V_T this would be turned off but here what happens is for this PMOS this gate voltage is 0 volts and this is the source here, the higher potential so this is the source which is at V_{DD}. So this one would always be on and so this output always going all the way up to V_{DD}.

In fact if you look at it this way that is suppose you have a V_{DD} here and the input is going from low to high, what you would expect if it is a perfect switch is the output would also go from low to high change like this. Now if you plot the resistance versus say V_{in} or V_{out} for the NMOS transistor, what is happening is as the voltage goes higher and higher, the gate to source voltage is reducing so the resistance of the NMOS go up like this. So I shall say R_{ON} for the NMOS and now for the PMOS what happens initially the voltage was zero here at the input.

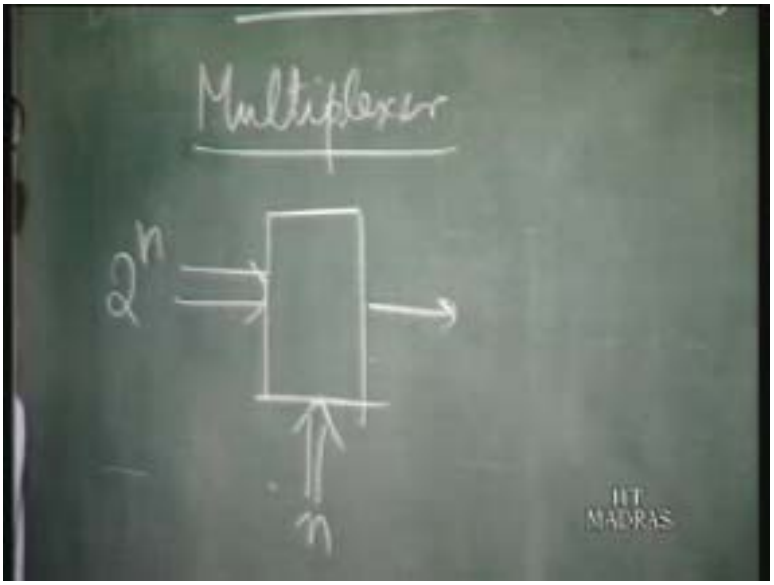
(Refer Slide Time: 34:25)



The PMOS was off because the gate to source voltage was zero but as this voltage is rising, the gate to source voltage is becoming more and more negative. Gate is zero, the source is becoming more and more positive, so the gate to source voltage is becoming more and more negative. So for the PMOS since the gate to source voltage is becoming more and more negative, the resistance is falling. We will plot for the PMOS, you have something like this R_{ON} for the PMOS.

The effective resistance which you have is actually the parallel combination of these 2 resistances. So the effective resistance you take the parallel combination is going to be something like this, I call this R effective say. So the effective resistance in a transmission gate is almost going to be a constant because of the parallel combination of the NMOS and PMOS. This is a more ideal switch compared to the pass transistor and the resistance is not going to alter depending on the input and output voltages. So this transmission gate may be used in place of a pass transistor and this is more compatible for CMOS technology as we can see because you have both PMOS and NMOS transistors.

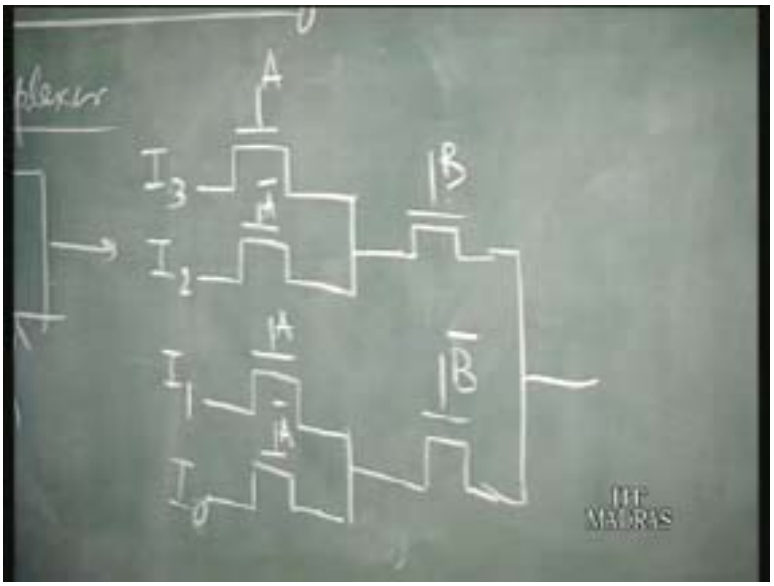
(Refer Slide Time: 36:14)



So when you are making a CMOS circuit, you can have this type of transmission gates. Basically we shall take up some pass transistor circuits to show that actually how you can realize complicated logic circuits using very few transistors. I shall draw the circuits using just an NMOS pass transistor but all this circuits can be realized using transmission gates by just replacing the pass transistor using transmission gates. The classic example or the use of this pass transistor logic is in the multiplexer circuit. Multiplexer has got, if you have n control inputs you have two to the power n inputs and one output.

So it is just like this, so this is you have n control input say and you have 2 to the power n inputs and 1 output and depending on the combination at the control inputs, one of the inputs is connected to the output. The output value you get corresponding to one of these inputs so that is a multiplexer. So how do you realize it using pass transistor? It is very simple. So normally if you use gates to realize that, you would require a large number of gates.

(Refer Slide Time: 38:18)



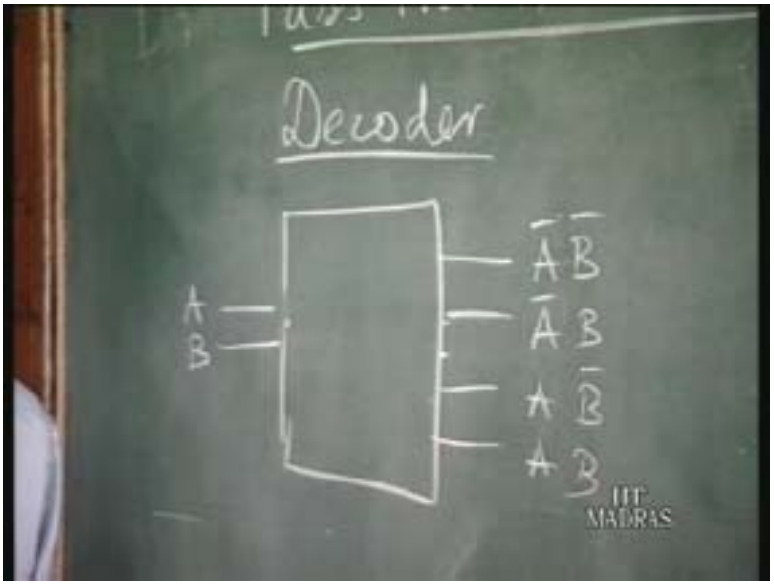
So suppose I want just say a 4 to 1 multiplexer that is you have 4 inputs and you require two control inputs to realize this 4 to 1 multiplexer. So let the control inputs be called A and B. Suppose this input is A, this is A bar, this is again A, this is A bar. This is B, this is B bar, so this is input 0, input 1, input 2, input 3. I think I should write it the other way around that is 0 1 2 3. When A is 0, B is 0, so 0 0 what happens is A bar is high, B bar is high. So this input here is connected to the output. You see that this line is connected, all other lines, all other inputs would be disconnected from the output. When you have zero one that is A is zero, B is one, then what happens is this goes connected that is when A is 0, this turns on and B is 1, it goes like this.

So for example when A and B are both high then this path is connected. You see that depending on the 4 possible combinations of A and B, one of the inputs is connected to the output and so you have a multiplexer using just 6 transistors like this. You would require many more transistor if you were to realize this using gates. For example if you take another circuit say a decoder circuit, you know what a decoder is. Again you have how many inputs? You have n inputs and you have 2 to the power n outputs and depending on the combination here one of these outputs is going to be high.

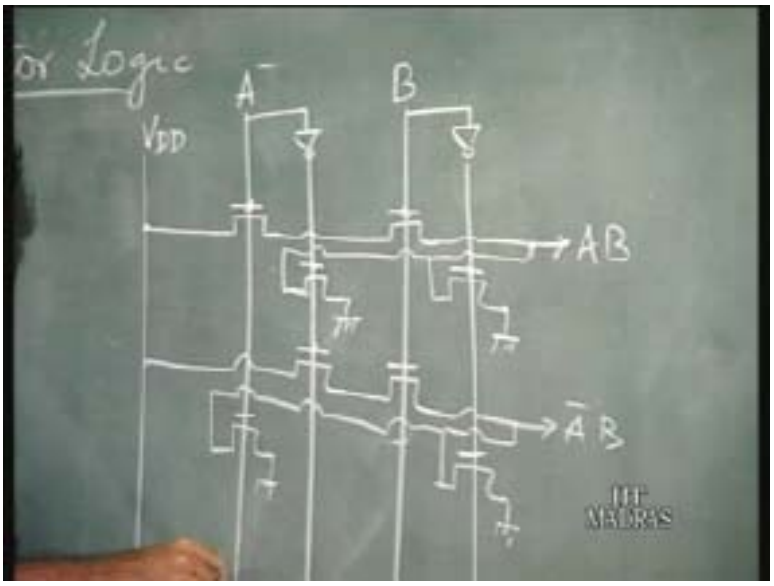
So if it is 0, just take the simplest case of two inputs and four outputs. For 0 0, one line

is going to go high. All the other output lines are going to be low. For 0 1 another line is going to go high like that. So how do you realize that using pass transistors? So basically what you have is in a decoder say if you have 2 inputs say A and B, you have 4 outputs, so you have A bar, B bar, A bar B, AB bar, AB the 4 possible combinations.

(Refer Slide Time: 39:59)



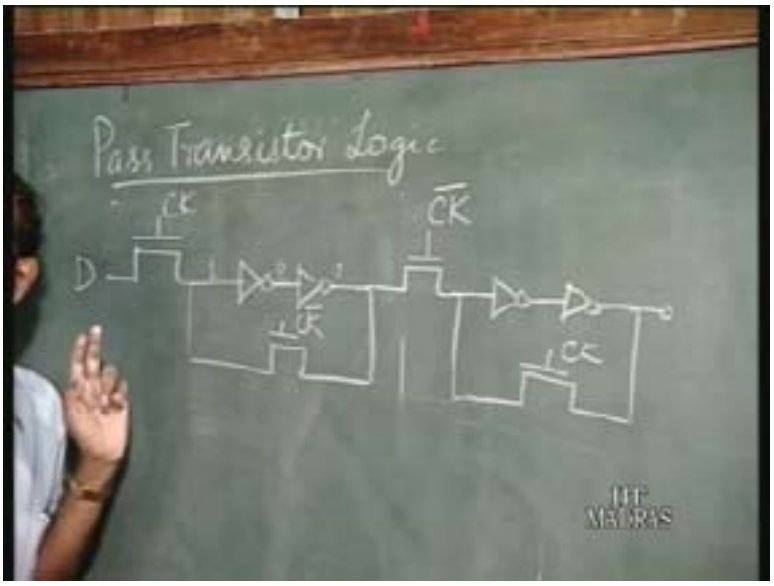
So depending on the input combination only one of them is going to be high, all other outputs are going to be low. So you can realize that using pass transistors, very simple configuration. You take $V_{DD.}$ then you have A, this is an A bar line, this is the B line, B bar line. So this is A, this is B so this output here is going to be high when A and B are both high. But again this is not complete because you have to ensure that when A and B or one of them is zero the output is undefined. So you have to ensure that whether A or B goes low, the output must go low. So you require two additional transistors here, you make this ground, this is connected to B bar. So when B goes low, B bar goes high, this transistor turns on. When this transistor turns on, this output must go low so this is connected here. So this output goes low.
(Refer Slide Time: 43:25)



Also when A goes low, this output must go low. So you have this connection here, these two transistors ensure that when A or B goes low, the output is going to go low. Only when A and B goes high, the output goes high so this gives you AB. Similarly you can have other combinations for example A bar B, so this is V_{DD} here. See if you have A bar B, you put like this. This is A bar B so again you must ensure that for the other cases output voltage goes to ground. So you have another set of two NMOS transistors here, so this is ground here, so this must be connected to the output. Of course you can draw the circuit for the other two combinations AB bar and A bar B bar.

The advantage here is we had already seen that in NMOS technology, you can have the gate made of poly silicon and when a poly silicon line overlaps a diffusion line, you have a transistor. So you see that how easy it is to make this circuit, it requires very little area because this line can be made of diffusion and these are the gates of the transistors, they can be running on poly silicon. So if you just run like this, wherever you don't want an overlap, you have to make this into metals say for example. So you can have a very dense circuit using a very little area. You can have a large decoder for example fabricated in a very small area using this particular type of circuit pass transistors. So that is the advantage of using pass transistor logic and also nowadays in VLSI large part of circuits are made out of pass transistors because of these advantages. Finally one more circuit I would just like to draw using pass transistors that is the D flip flop. You can easily make a D flip flop.

(Refer Slide Time: 46:57)



D flip flop is just a D flip flop with which is enabled when the clock is high will pass the input to the output when the clock is high and when the clock is low, the previous output is going to be retained. That is very easily realized using a pass transistor. This is the D input and say this is the clock here and you have two inverters and here you put clock bar.

So here when the clock goes high what happens is whatever is the D input is passed on here and you have two inverters here, so basically the output of this two inverters is going to be the same as the input here. So at the output you have the same as the D input but when the clock goes low, you have to retain the previous value. So this pass transistor cuts off, this transistor turns on so you have two inverters connected like this in series (Refer Slide Time: 45:47).

So whatever is the input, it is fed back. So if this is 1, this becomes 0, this is 1 so it's a perfectly stable system and it is going to retain whatever is stored here. So this is a D flip flop. Now if you want sort of master slave operation that is you want the output to change only at a particular instant and not when the clock is high, what you have to do is you have to have a second stage and here what you do is you make this clock bar and this input here is clocked.

This is going to give master slave operation because this slave flip flop operates only when clock is low. So what happens is when the clock is high, whatever is the D input is it can go up to here but it cannot pass through. So what is going to happen? So this is retained here now when the clock goes from high to low this turns off, this turns on; whatever is retained here is passed on to the output but after that instant this clock is off. So the input will have no further effect on the output and whatever was here is going to be retained at the output of the flip flop (Refer Slide Time: 47:15).

So only at the instant when the clock goes from high to low, the output is going to change and after that there can be no further change because this master is going to be off when the clock is low. So this behaves as a master slave flip flop, D type flip flop. So you see again how easily you can realize this type of logic circuits using pass transistor logic and that is why it is so popular because one hand you save space or area using pass transistor logic and also the power dissipation is very much reduced number one because you are saving on a lot of circuitry and also because pass transistor by themselves do not consume any power. They are just transferring the input to the output.

So that way you save a lot of power so both of which are required in a large scale integrated circuit. So with that we come to the end of this section on pass transistor logic. In the next class we shall take up another variation of CMOS which has become quite popular nowadays is bi CMOS where actually you use bi polar transistors together with CMOS to achieve logic circuit. The advantage of using bi polar transistors is that you can increase the speed of operation at the same time, you retain all the advantages of the CMOS in the bi CMOS. That is the low power dissipation that is zero static power dissipation. So the problem was initially that it was difficult to fabricate both CMOS and bi polar transistors on the same way because the requirements where different but once that problem has been sorted out, one can really fabricate bi polar transistors and CMOS on the same wafer. One can really derive the advantages of both CMOS and bi polar transistors, the drive capability of bi polar transistors in one logic circuit. So that is superior to CMOS in many ways. We shall take up that for discussion in the next class and see some applications of bi CMOS logic.