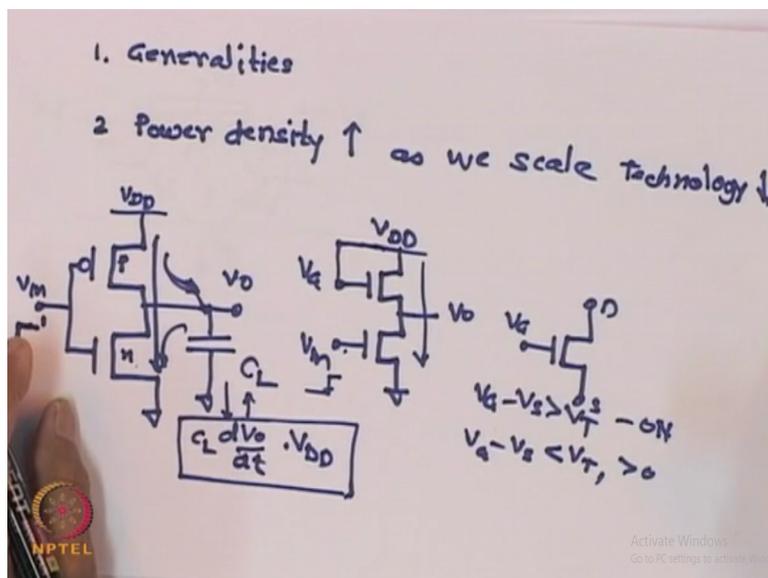


**Advanced VLSI Design**  
**Prof. A. N. Chandorkar**  
**Department of Electrical Engineering**  
**Indian Institute of Technology – Bombay**

**Lecture - 07**  
**Power Estimation and Control in CMOS VLSI Circuits (Contd.)**

We will continue with our talks about and lectures about power estimation and control in CMOS VLSI circuits and I repeat I am from IIT Bombay. Actually before I go to this, what we did in our first talk was the following.

**(Refer Slide Time: 00:40)**

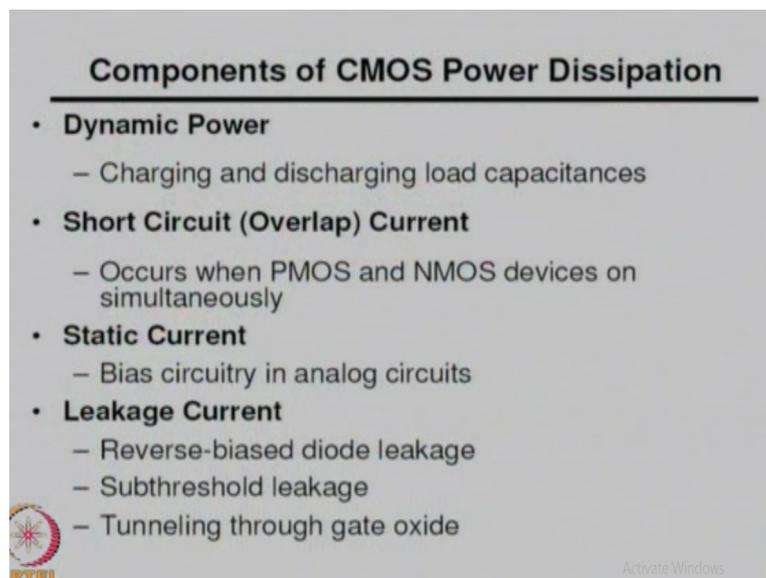


We talked about generalities in which we showed the importance of power in case of CMOS and other MOS circuits. We just discussed that systems like say mobile or many other systems do require low power circuits and we also have circuits which requires high performance that is high speed circuits. But in the either case, one has to accept that the power reduction is always beneficial for variety of reasons as long as you do not lose on the given performance.

So then generalities we talked about many of these things. Particularly we talked about the power density increase as we scale down was increasing due to as we scale technology node. As technology node goes down from say 90 nanometers down or 0.13 micron down to 22 nanometer and below, one of the worries right now to all of us is to increase in power density. So in any case whether it is high performance or low power or low standby.

One has to reduce the power dissipation in the circuit. For other reasons that the package which can remove the heat probably has to be now must get larger attention because at the end of the day, the cooler the chip better is the performance. Today we shall continue with our system and we move now to our today's lecture and we will start with where the power dissipation is components of CMOS power dissipation.

**(Refer Slide Time: 02:24)**



Okay. So we will start with the CMOS dissipation today and we will continue with this in my next lecture as well particularly the effort to reduce the power dissipation. The four kinds of power are visible are written here which actually consumes power in the circuit which then uses CMOS circuits. First one is of course is the dynamic power, charging and discharging load capacitances, this of course is, we will soon that basically dynamic power is nothing.

But a switching power, that means when the input of a CMOS circuit changes from 0 to 1 or 1 to 0 the output capacitances charges or discharges at the load capacitance and charging or discharging current which is nothing but the displacement current due to the capacitor  $cdv$  by  $dt$  that multiplied by the voltage essentially gives the power which is consumed from the battery.

And how much is stored in the capacitance and how much is lost in the capacitance afterwards during the other transitions is what we have discussed in the case of dynamic power. The second and one of the other important current which is worrying these days more often than earlier times is when the transition occurs from 0 to 1 or 1 to 0, we know in case of 0 to 1 inputs, initially our n channel CMOS transistor.

If you see the figure again, if you see the typical CMOS circuit which is shown here and this is my output load and this is my output and this is my input and let us say I am going from this transition to 0 to 1, at that time initially the n channel transistor was off and p channel was fully on charging this capacitor to VDD, however when the transition occurs, this transistor also turns on and this transistor starts from saturation.

It has to go to non-saturation mode. This breaks up into saturation and goes to non saturation mode. So at times when both transistors are on, there is a connection between VDD and ground and this power dissipation essentially we call it short circuit current, power dissipation due to short circuit current. That means both n channel, p channel are turned on, their status maybe saturated or non saturated depends on the input.

However, they will continue to till the transition is completed. Once the input transition is over, then there is no current really from the VDD to this except for the dynamic current which is due to the capacitor charging or discharging. So essentially we are saying that short circuit current will always occur in any CMOS circuit because of the transition itself. So one can understand from a very simple thinking that if the transitions do not occur.

Obviously there will be no short circuit currents. So there is no dynamic power. So one of the method of reducing the power dissipation in circuit is to see that your data is such that it gives lesser transitions compared to the data which has larger transitions. Now this of course how to code it, is another area where low power circuits can be designed. The third possibility of current flowing in a CMOS circuit is bias circuit.

And analog circuit particularly which is not so much in digital, or it does occur in even digital when we have an NMOS inverter which is rarely used but could be used. We have  $V_g$  here, we have  $V_n$  here, we have output here, we have VDD here, this is my output, this of course  $V_g$  can be larger than  $V_t$  of this n-channel transistors so that it turns on, but if  $V_g - V_t$  is less than  $V_0$  VDD.

We say it is in non saturation and if we connect it to something like this, the n-channel load will always be in saturation. So this is something NMOS inverter and in NMOS inverter, except for when the  $V_m$  is slightly lower than  $V_t$  when both devices are off. For all other case

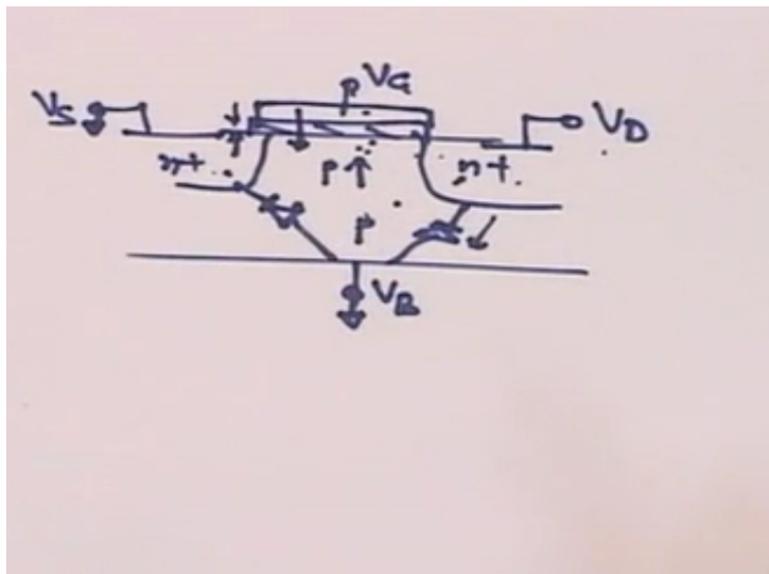
we can see a constant current can flow and that is called the static current in NMOS circuits. However, in CMOS this situation does not occur.

And therefore we can say always CMOS circuits are low power circuits compared to NMOS circuits. The fourth and the most worrisome current mechanism which is flowing in a CMOS circuit particularly which is worrying us most in a less than 90 nanometer nodes of technology circuits, this is called a leakage current. What we assume in a normal transistor, if you have this transistor and let us say you are applying  $V_G$ , this is your s, this is your drain.

As long as we say  $V_G - V_S$  is greater than  $V_T$  we say transistor is turned on. So essentially what we say is on, but we can also think of a case when  $V_G - V_T$  is less than  $V_T$  but greater than 0. Under this case one assumes that device should have been switched off and in real life this does not occur. This is essentially means, this occurs as we have last time discussed that because the device is in sub threshold region and in sub threshold region.

There is a current flowing even if  $V_{GS}$  is less than  $V_T$ . Now this sub threshold current is essentially we shall discuss later, it depends on the kinds of technology we use and at the temperature at which device operates and it increases in both cases. Now sub threshold leakage is one of the major worries. But the other leakage currents are always present in a MOS transistor.

**(Refer Slide Time: 08:35)**



If you look at a typical MOS transistor, one can see that a MOS transistor there is a source, let us say it is an n channel transistor and there is a p substrate and this is your gate oxide and

this is gate oxide and this is your gate contact. This is your drain and this is your source and this is your bulk. So we say even if we ground the bulk and if we apply  $V_D$ , drain positive voltage for n channel device at  $V_{DS}$  and if source is also grounded.

You can see there is an NP diode here, this is N + V diode here and this diode if you see clearly even if at 0 bias this device is reverse biased and there is a leakage current in this device which maybe essentially close to generation currents, whereas in the case of heavily reversed biased diode there is a leakage path which is essentially the reverse bias diode leakage.

So whatever technology one uses these diode currents are going to be always present and at lower technology node the p may actually, p looping may increase and in which case the leakage current will be further higher in the case of lower technology devices, lower technology node devices because lower technology node devices the substrate doping has to be increased to keep the  $V_T$  within your check.

And that actually enhances the reverse biased diode leakage and one can say that this diode current is essentially  $e$  to the power  $qV$  by  $n kT - 1$  and this current is proportional to temperature, larger the temperature larger the current and therefore larger the power dissipation in chip, this current also enhances over the time. So this issue of leakage current is not trivial and the third possibility as we shall show figure little later.

This since everything is scaling this oxide insulator thickness is also scaling and if that is scaling below less than say 20 or 30 Armstrongs which may be required for a 22 or 32 nanometer nodes, then this gate oxide and since voltages are normally not scaling, the fields across these insulators are very high and prompting some carriers to tunnel through from the gate to the substrate or substrate to the gate and these are called tunneling through gate oxide.

There is of course additional current which is called gate induced leakage which we shall see little later. So we can see there are four kinds of power which we are worried about, first is the dynamic power which as I say, if you look at this dynamic power, it is essentially charging and discharging of the load capacitors. So essentially current in this or current going out of this is  $CL dV_0$  by  $dt$ .

And if you multiply it by  $V_{DD}$  this is the power dissipation which is dynamic. Since it is time dependent it is called dynamic power dissipation. In the short circuit as I said, except when the transition occurring at the input, the current flows from  $V_{DD}$  to the ground and that current we call short circuit current and the third of course as I say static current which I keep saying is much smaller in the case of CMOS.

There is a possibility of some current in the case, large static current in the case of NMOS. And the fourth of course I say diode leakage plus the sub threshold leakage plus the tunneling oxide, they all constitute the leakage current. So if I want to reduce the power dissipation obviously I must control all of them and try to see how much of this currents can be minimized at the technology nodes of our choice and that will decide the power dissipation.

**(Refer Slide Time: 12:35)**

**Dynamic Power Consumption**

- Define effective capacitance  $C_{eff}$ :

$$C_{eff} = \alpha C_0$$

$$P_{switching\_inverter} = \frac{1}{2} C_{eff} V_{DD}^2 f_{CLK}$$

- To minimize switching power
  - Reduce  $V_{DD}$
  - Reduce  $C_{eff}$

The diagram shows a CMOS inverter circuit with a load capacitor  $C_L = C_{eff}$  connected to the output node.

This is what I said already, so here is a figure which shows dynamic power consumption. We define an effective capacitance as alpha times  $C_0$ . Now this  $C_0$  is the output capacitance at this node, but we have now introduced another term which is called alpha. Now this alpha as we shall go ahead and look into this value, this is essentially called the activity coefficient.

So what it means that depends on the input change whether it is 0 to 1 and 1 to 0, no change when the data stream appears on this. There will be transition in this or there will not be any transition here, so alpha essentially means how much times the capacitor in a cycle charges or discharges, that how many times were it taken care through a factor called alpha which is called activity coefficients.

Now if you see activity coefficients, we take care of that through the load capacitance, we called effective load capacitance which is the net capacitance, which we can define as the output capacitance time the alpha. Please remember alpha is not a constant quantity, it varies with the data and therefore C effective also varies with the data. So for the switching power or the dynamic power as it can be called, essentially can be figured out.

And we will derive this expression in a short while that see power dissipation due to dynamic power consumption is half C effective square VDD multiplied by fCLK where fCLK is the clock frequency at which the data is changing or 1 upon fCLK is the period for which this input is actually retained or changed. Now, to reduce the dynamic power consumption this expression which we wrote, half c effective VDD, of course one cannot reduce fCLK.

Because ultimate aim of any circuit these days is essentially the increase of top frequency, larger speeds is what is desired. Therefore, this essentially is going to increase as the technology is improving. So we are trying to go for higher and higher frequency gigahertz and few gigahertz to higher gigahertz. However, if you wish to reduce switching power then there are two terms which you must take care of, one of course is the power supply.

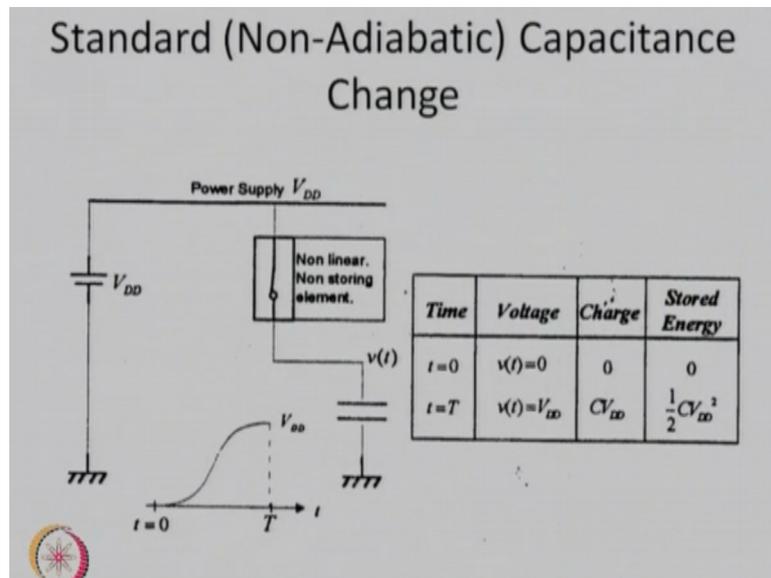
If you reduce VDD, then of course the square term and therefore switching power will immediately go down as it is proportional directly to the square of VDD. However, we know in scaling of technology, everything else can be scaled but it is not a constant voltage scaling. It is not a constant field scaling, it is a constant voltage scaling and voltage does not scale as much as the other dimensions.

And since VDD, it cannot be scaled down very much because of the noise considerations this VDD term is also not, you cannot reduce very much because otherwise you will not be able to have enough currents also. Now if you look at the third term which is C effective, so obviously one can reduce the power dissipation by using switching power dissipation, if I say that I can reduce C effective.

And essentially C effective also cannot be reduced directly as you see because C0 is something fixed for the circuit, however what we can change by variety of ways is to change this alpha should go down and if alpha goes down then C effective goes down and if C effective goes down then the next switching power will also go down. So a typical CMOS

inverter, the dynamic power can be controlled or can be minimized by reducing the VDD and reducing the activity coefficient alpha.

**(Refer Slide Time: 16:18)**



Now of course this word essentially what I wrote will be clear after, the end of this I will talk about other kind of energy dissipation, so other kinds of circuits which are adiabatic circuits. This circuits which I am showing is called adiabatic. Non adiabatic circuits are those according to law of thermodynamics, the adiabatic changes actually will always consume energy irrespective because of law of conservation of energy.

Whereas in the case of non adiabatic, in the case of adiabatic probably you know some thermodynamics. What we are doing is, we are reusing the energy stored once and therefore one can say as if the net loss of power is much lower or you can say theoretically saying it can be even 0 because we are using the energy ahead and therefore you are never consuming or dissipating it.

However, such adiabatic circuits were tried hard in the many years to go and including this myself I have worked earlier with some few students, however, there are many issues in making adiabatic circuits. They are better in as far as power is concerned but they are not as good in speeds. However, in the case of DSP based circuits, some of those certainly can be implemented using adiabatic circuits.

So coming back to non adiabatic change when the power is going to dissipate, here is a circuit shown here. This is some kind of a nonlinear, non storing element. This is like shown

as a switch. This switch is essentially what we replicate it by MOS Transistor or any other transistor and this is of course I keep saying it is a non linear and non storing element.

And whatever capacitances associated with this are clubbed apart from the external capacitance into a known external capacitor  $C$  plus the capacitances due to whatever non linear element used is  $L$ . Now the way the transition occurs, let us say I want to charge this capacitors, so obviously for VDD I switch on the switch. So I have a path from VDD to the capacitors.

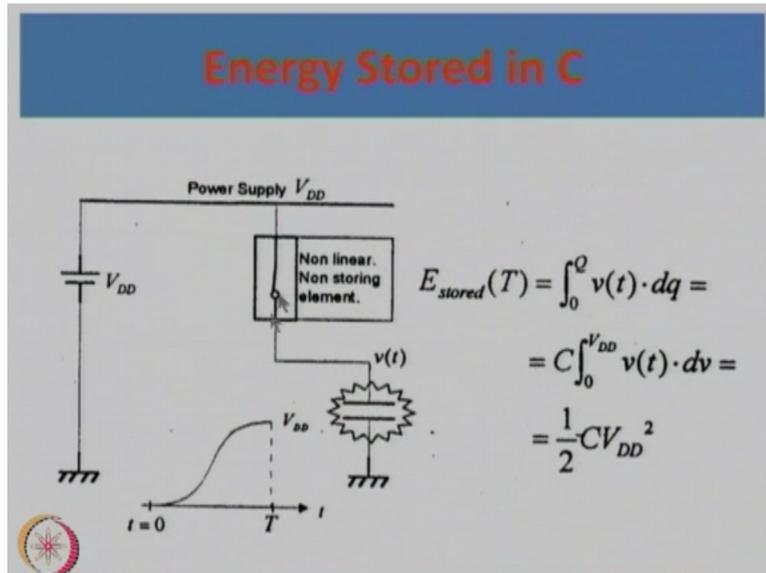
So this current will flow through this and as this current pass it will start charging this capacitor and the transient is shown here as  $t$  is equal to 0, there is no voltage on this capacitance and as you increase, this is  $V_T$  graph, so as you increase time and the switch is on, you can see from here capacitance start charging and it reaches, the voltage maximum reaches to the VDD, that mean whole VDD then will appear across the capacitor.

Table wise we can say, in time at  $t$  is equal to 0,  $V_t$  at this capacitor is 0. There is no charge stored in the capacitor and there is no stored energy anywhere. Please remember capacitor is a storage element of the charge and therefore energy. Please remember  $Q$  multiplied by  $V$  is capacitor but  $Q$  multiplied by  $V$  is energy and therefore any  $Q$  is equal to  $CV$ , so essentially we are saying  $CV$  square will be energy term.

So we can see from here, this figure  $t$  is equal to 0,  $v_t$  is 0 and no energy storage, no charge on the capacitor. However, as the transient goes from 0 to time  $t$  which is the period the capacitor charges to full VDD at  $t$  is equal to  $T$ , therefore  $V_t$  is VDD and the charge is  $Q$  is equal to  $CV$ ,  $C$  is equal to this capacitor multiplied by VDD and the energy stored on the capacitor as we can easily find out by integrating this is half  $CVDD$  square.

So this is how the transient will actually lead to. So you have a, when your output is charging from 0 to one, essentially you are storing an energy of  $CVDD$  square.

**(Refer Slide Time: 20:19)**



Okay, the word which I said can be derived mathematically as well. This is the same figure again. This is my power supply, this is the switch which is turned on at  $t$  is equal to 0 and please remember I repeatedly saying it is non linear because most transistors which will use as a switch have a nonlinear characteristics, but they do not store and the storing element part of that has been taken care in this net output capacitance.

So initially if you want to know how much is energy stored over a time  $t$  then we know the method is  $QV$  is essentially the energy and stored from the charge stored going from let us say total charge at the end of  $t$  is equal to  $T$  is  $Q$ , then we say energy stored in time  $t$  is 0 to  $Q$   $v t dq$  but we know  $q$  is equal to  $CV$ . So if I substitute  $dq$  is equal to  $Cdv$ , say if I get this  $dq$  replaced by  $C$ .

Then the voltage goes from further getting from 0 to  $q$ , is voltage goes from 0 to  $V_{DD}$ , so we integrate it from 0 to  $V_{DD}$  times  $v t dv$  and if I integrate this simply then I becomes half  $CV_{DD}$  square. So energy stored by connecting the power supply to the capacitor through a switch in time  $t$  stores an energy which is equal to half  $CV_{DD}$  square, however this is not the power supplied by the battery.

**(Refer Slide Time: 21:39)**

## Energy Supplied by the Battery

$$E_{Sup}(T) = V_{DD} \cdot Q =$$

$$= V_{DD} \cdot CV_{DD} =$$

$$= CV_{DD}^2$$

Let us look at the battery power, battery is here which has a voltage of VDD. So we want to see what is the supplied power in given time. So voltage is always full going to up to this and the net charge at this point when it reaches full charge, then the energy which actually power supply is going to give you is VDD times Q, but Q is essentially the capacitance multiplied by VDD, so which is CVDD square.

We have said that the energy supplied by the power supply is CVDD square where as power stored in the capacitor is half CVDD square when the transient goes from 0 to 1 at the output. So one can see that supply is almost double that what is being stored which essentially is trying to see where the rest of the power is going.

**(Refer Slide Time: 22:32)**

## Heat = Energy Dissipated (Non-Adiabatic)

$$E_{Dis}(T) = E_{Sup} - E_{Stored} =$$

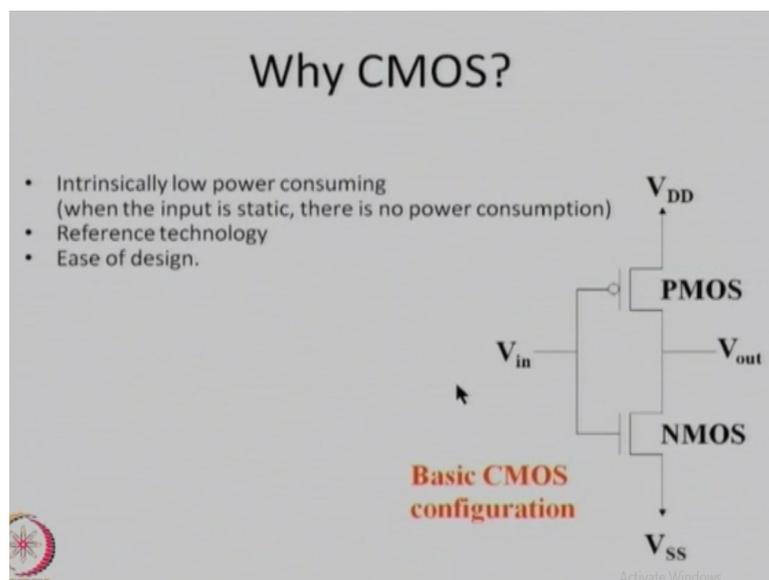
$$= C \cdot V_{DD}^2 - \frac{1}{2} CV_{DD}^2 =$$

$$= \frac{1}{2} CV_{DD}^2$$

So obviously this energy is getting dissipated. You are actually supplying CVDD square, you are storing only half CVDD square which means rest of it , half CVDD square energy is dissipated in this. Now where does this dissipation can take place. Capacitance is essentially we call non dissipating element as far as the heat is concerned. I am not saying there is no ionization loss but practically that loss is much smaller.

So obviously the power which is being dissipated, if it is not here obviously it must be here because if the power supply is giving CVDD square, half CVDD square coming here, the non linear non storing element must dissipate a power of half CVDD square and as I already said that this non linear non storing element is nothing but going to be your NMOS or any other transistor switch, the power is dissipated in your device which is your acting as a switch which is equal to half CVDD square.

**(Refer Slide Time: 23:38)**



Now many queries are asked that why CMOS? So we have already see few minutes ago that in NMOS there is always a static current flowing when the transistor is, the n channel transistor is, driver transistor is on and for all the rest of the transistor it continues to conduct the current. So since NMOS will always be high power dissipating circuit, full NMOS circuit, therefore we look for circuit which are intrinsically low power, dissipating circuits.

So it is found that since input in the case of input static there is no power consumption. If there is no transient, the VDD to ground path is always broken in the case of CMOS when  $V_{in}$  is 0, PMOS is conducting, but NMOS is switched off, therefore no current, at least no

leakage current exist, but no dynamic current or no static current flows, if  $V_{in}$  is 1 then this device turns off, this device turns on.

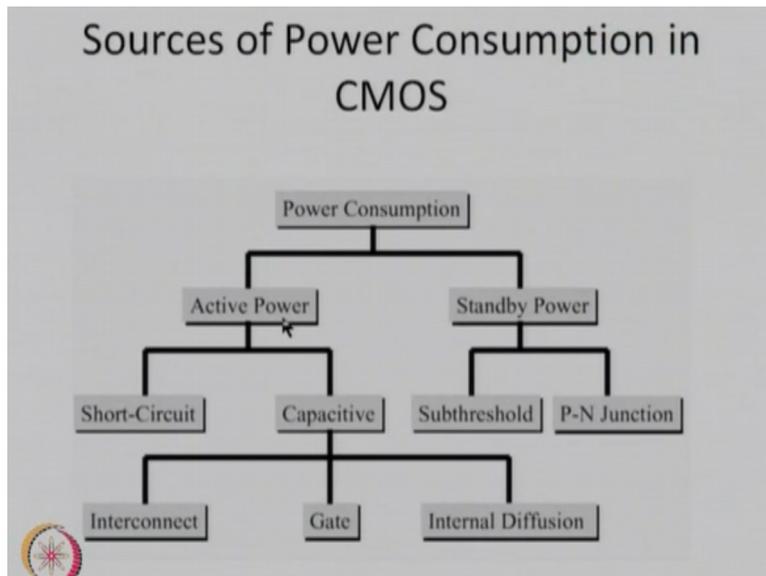
But since there is no connection between VDD to VSS in either case. So the input is either 0 or one, there is no power dissipation as far as switching is concerned, as our static current is concerned. However, we know when the transient occurs this status will change and then there will be a short circuit current. Now why CMOS is the other reason is that the world over last 30 years have switched over from NMOS to CMOS majority of course was because of the power.

And it has become a standard technology for almost all VLSI systems in the market barring exceptional there are some bi CMOS circuits, there are also bipolar circuits also in the market for very strategic application where speed is the only criteria that is high performance circuits. Other than this for almost every circuits you are actually looking into many systems which you use, everyone is using CMOS.

What we probably are changing is the material possibility. Instead of silicon we may use something, instead of silicon dioxide we may use high-Ks. In the case of instead of silicon, we may try germanium, we may try gallium arsenide or a host of 3 to 5 to 6 compounds. We can modify the structure of gate, PMOS or NMOS by making it thin fed, but in either case the basic circuit component will remain PMOS and NMOS as a complimentary.

And this basic block will remain constant for this. Now having worked 30 years, almost every one of us have now realized how to design circuit using CMOS and therefore the ease of design, the availability of cells already or IPs available in CMOS, everyone will still continue to use CMOS for next 10 to 15 years, take a word from me.

**(Refer Slide Time: 26:31)**

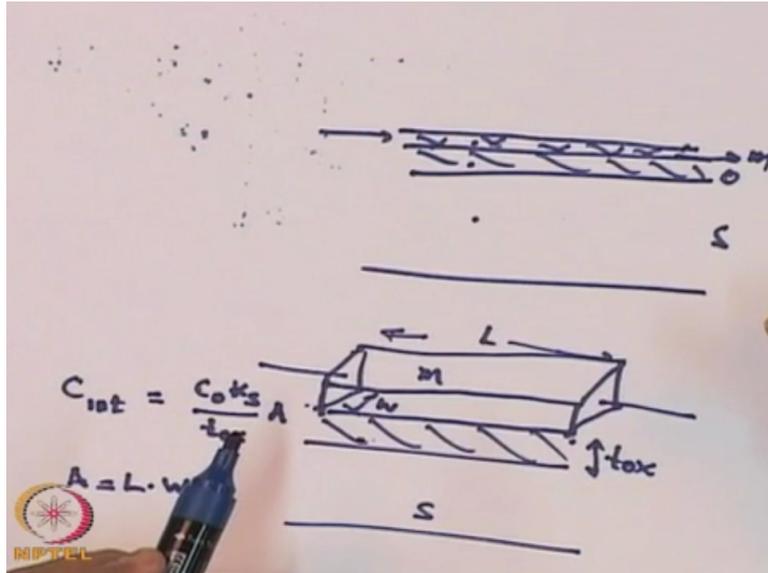


Now to summarize before we go ahead in dual power, let me tell you what are the power consumption in CMOS. Net power consumption is divided into two parts, one we call as active power and the other is we call standby power. The active power essentially has two parts, one is that so called dynamic power which is capacitor charge discharged which is essentially what we say switching power.

Another of course when the transients goes through both N channel and P channel turns on for a while, during that transient and we have a short circuit current. The capacity current also stand from the variety of regions of a CMOS circuit for example, the capacitance contribution comes strongly these days from the interconnects. Interconnects are of course are the metal lines which are running on an insulator.

For example, if this is your silicon, I may have insulating field oxide on the top and the top of this my interconnect line, metal line may run. Please remember the current is or the signal is going in this metal line, but it does see metal oxides semi conductor.

**(Refer Slide Time: 27:45)**



So you have an MOS capacitor sitting for this and the capacitor value will be decided by the length and the thickness and the width of this metal line. For example, if you see three dimensional picture, this is your metal line, this is your metal thickness, now the resistance is decided by  $\rho L$  by  $A$  of this. The current maybe flowing on this. But there is an insulator below here.

There is a semiconductor below here. This is your metal. So depending on the length of this and also the width of the metal line, this is the area of the metal line sitting on the oxide, so a MOS capacitor which we call C interconnect will be  $\epsilon_0$  multiplied by  $K_s$ ,  $K_s$  is that of whatever semiconductor used multiplied upon the oxide thickness or the insulator thickness, which ever insulator are used multiplied by area of the metal which area here is  $L$  multiplied by  $w$ .

So one can see from here that larger the  $L$  and  $W$ , longer the length of interconnect are used, larger capacitance will be actually accruing at the note of the output and therefore large capacity effect can be seen and hence, larger the capacitance, we already said larger the capacitance and larger is the active power dissipation. The other of course is a gate, gate has a capacitor sitting there  $C_0 K_s$ .

And therefore that capacitor is also contributing to the net capacitance and finally as I said there are other velocity capacitance in most transistors which are essentially due to the diodes there and those are called internal diffusion capacitance. So drain to bulk, source to bulk are the capacitances or gate to drain and gate to source are the other velocity capacitances which

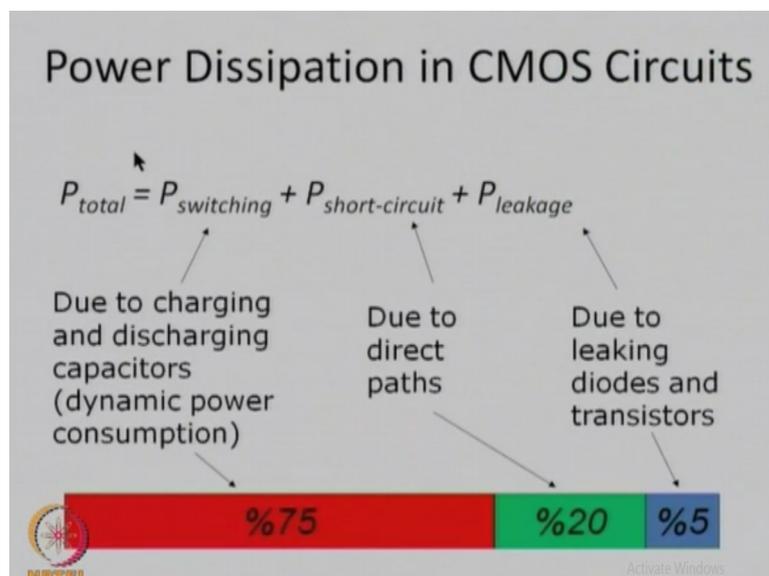
also will be attracted, will also be connected at the output node or any capacity of the output node and therefore add to the net capacitance.

Larger the capacitance, larger is the consumption. If you look at the other side which we call as standby power, we have two source of power dissipation with standby, standby essentially circuit is on, but not acting. What does that mean like, a mobile you keep it on a standby mode, whenever you actually send them, use, talk on a phone or a mobile phone or you receive a phone any call at that time the circuit is switched on fully.

And the power is as much as active power. However, even if you are not using the phone for any such purpose, the device remains on for a while because otherwise you cannot bring the on state faster than when someone calls you, or you want to call. So this has to be kept in standby mode and that standby power is essentially because the two leakage, main, of course there are multiple leakage as we may see later.

But there is a sub threshold current and the diode leakage current which actually keep constantly leaking the current and the battery has to keep supplying that power. This is called the standby power.

**(Refer Slide Time: 31:03)**



So in a normal CMOS circuit, I would say there are three components. The total power dissipation can be due to switching, due to short circuit and due to leakage. And in our older technology and that is why I brought this numbers, please remember, I repeat switching

power is essentially charging, discharging of capacitor which is essentially dynamic power consumption.

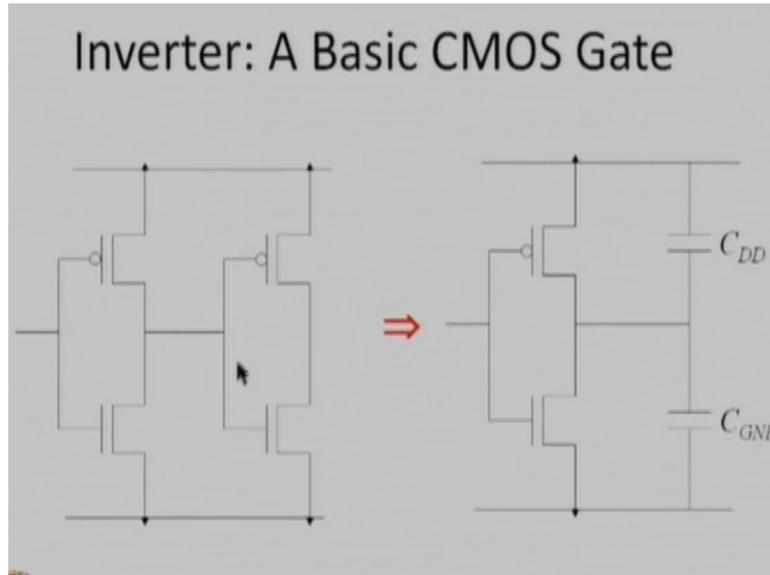
Then this half CV square end of time for which it switches, we shall see that. The P short-circuit is essentially due to the direct path between power supply and down when the transient goes through the input of a CMOS circuit and finally there is a leakage due to leaking diodes and transistors. But you can see till very late, till as 95 to 98 kind of times, the technology was just we are going below 0.13 or around that time we were still above 100 nanometers.

The most of the power once sees is essentially the dynamic power which is 75%. The short circuit data dependent power dissipation is 20% and finally there is a diode leakage current which was 5% and actually it is here we shall see later when the technology node is going down from 90 nanometer to 22 or below, this power actually may overpower both of these and our worry starts then because if this is larger than the sum of these.

Right now this is 95% and this is 5%, it has been observed that this may become 66% and this two together may give 33% which is very bad because essentially it means even if you are in a standby mode you will consume much more power and the effort has to be reduction in leakage powers. However, till such time we come to that, we will discuss right now as if to reduce dynamic power we must work out.

Because this is 75% power which is going in the dynamic power or switching power. Of course we also should look for how good we can reduce this 20% power, to many a times this power reduction is not so very easy as we shall see later.

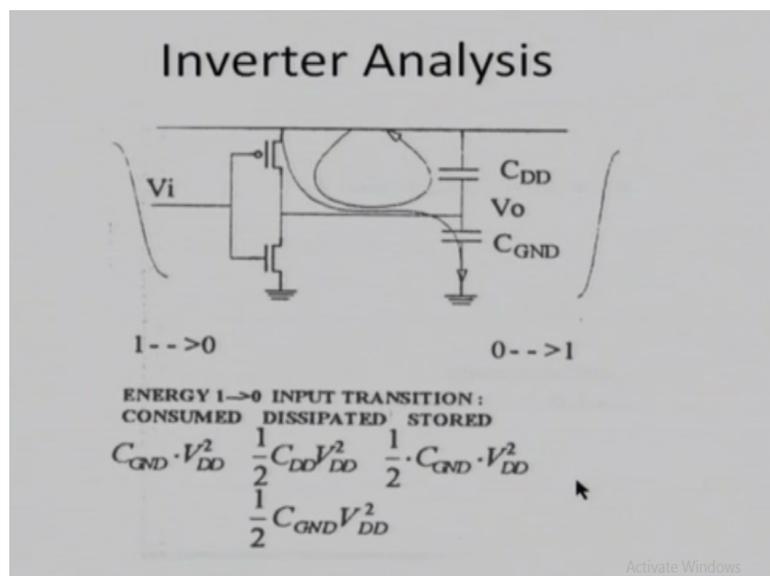
**(Refer Slide Time: 33:22)**



So a typical basic CMOS gate is shown here. A CMOS gate is driving another CMOS gate which are same size,  $W$  by  $L$  for both n channel, p channel and can be circuit wise made equivalent of a capacitance at the output node is equal to  $C_{DD}$  as capacitance bulged at the output node equal to  $C_{DD}$  to the power supply and other capacitance is the  $C_{GND}$  which is to capacitance going.

This is equivalent of net capacitance which includes velocity and every other capacitance at this node because of this, because of the bulk to this, because of this plus the capacitance due to this, at this node, and the interconnect, all these capacitances are can be taken care by two such capacitance, we say one going from output node to the power supply capacitance, the other is going from output node to the ground which we call  $C_{GND}$ .

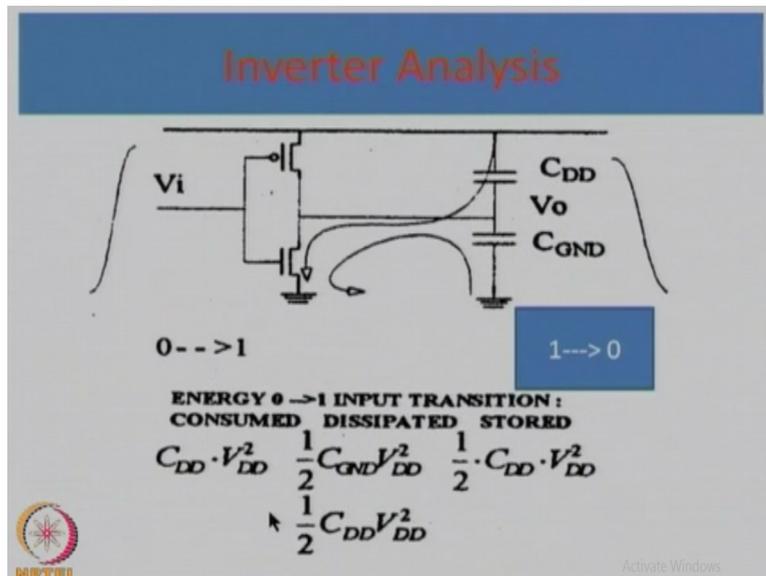
**(Refer Slide Time: 34:20)**



So if you do an inverter analysis as shown here, if we are doing a transition from 1 to 0 as shown here and the output going from 0 to 1 due to the inverter action, one can see there are 3 energy consumed, energy dissipated and energy stored, so power supply is going to give  $V_{DD}$  square as the power supply due to charging, then we can say half  $C_{VDD}$  square will be stored, this dissipated in half  $C_{GND}$  square would be essentially stored on this.

This other power is lost on this whereas if you have a discharge transient, this is the only power. So we say dissipation is only this much, that is half when the output is going on 0 to one, so this is the only loop which we can see, so we have the power dissipation of  $C_{GND}$  square.

**(Refer Slide Time: 35:17)**



Now continuing with this, if the transition is opposite,  $C_{VDD}$  0 to 1 and this is 1 to 0, but the numbers do not change because essentially we are saying there it was  $C_{GND}$ , here it becomes  $C_{DD}$  and in most cases these two capacitances are same by design and therefore one can see, even if you call it  $C_{DD}$  or  $C_{GND}$  it does not matter, so we call  $C_{VDD}$  square half, half  $C_{GND}$  square half  $C_{DD}$   $V_{DD}$  square.

So essentially if the transistor are identical and matched to  $W$  by  $L$  so that the mobility difference is taken care, mobility ratio is taken care irrespective of what we do, whether you make a transition from 0 to 1 or 1 to 0 at the output or the input, the power given by the power supply is  $C_{VDD}$  square and power consumed by the devices half  $C_{VDD}$  square and power stored on the capacitances is half  $C_{VDD}$  square.

**(Refer Slide Time: 36:22)**

**Energy Consumed  
(related to Battery Power)**

Energy consumed due to a complete cycle 0→1→0.

$$E_{\text{cycle}} = E_{01} + E_{10} = C_{GND}V_{DD}^2 + C_{DD}V_{DD}^2 = (C_{GND} + C_{DD})V_{DD}^2$$
$$C_0 = C_{DD} + C_{GND}$$
$$E_{\text{cycle}} = C_0V_{DD}^2$$

 Activate Windows

So we have just talked about energy consumed due to complete cycle going from say 0 to 1 and 1 to 0, this is the complete transition. So in a one cycle time, we have one transition power energy dissipation due to 0 to 1 and other dissipation is due to 1 to 0 and we add these CV square terms, C ground + CDD multiplied by VDD square.

And typically this is what the C0 term which is the output capacitance at that node which is nothing but some of CDD + C ground and one can also see from here that this C0 multiplied by the activity coefficient, how many times the input changes 0 to 1 or one, output changes 1 to 0 or vice versa, we call one cycle time, in one cycle time the power essentially is C0 VDD square.

Just remember we are talking of one cycle time. But if the data is coming in number of cycles what is the power dissipation or e-cycle available to us.

**(Refer Slide Time: 37:32)**

## Power-Delay Product

$$PDP = P_{av} t_{pd}$$

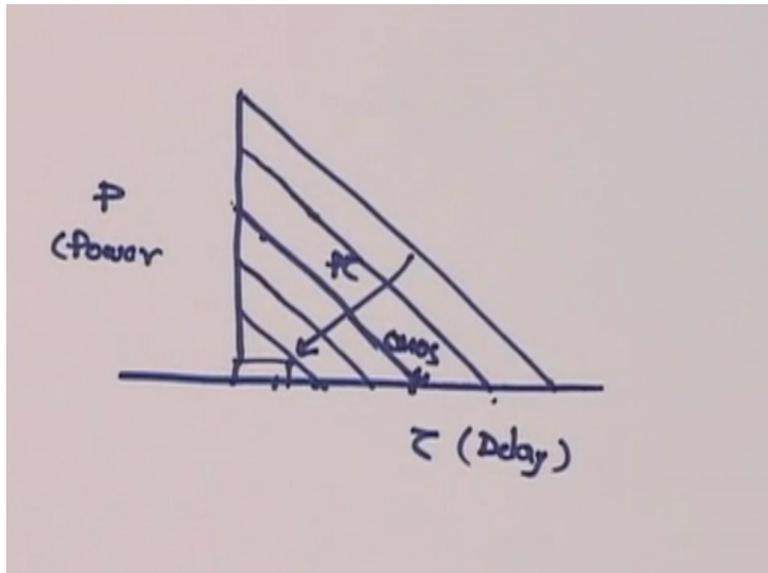
- Product of average power and propagation delay  $t_{pd}$  is generally a constant
  - For given technology and fixed gate topology
- Define propagation delay as average between low-to-high transition delay and high-to-low transition delay:

$$t_{pd} = \frac{t_{pLH} + t_{pHL}}{2}$$

- PDP = Energy consumed by gate per switching event (Watts x seconds = Joules)

Now the next term for all power people is essentially called power delay products which is we say that average power multiplied by the average propagation delay TPD, the product of this two is called the PDP and it is normally for a given gate topology and for a given technology this may remain constant.

**(Refer Slide Time: 37:56)**



This is very important because we actually PDP is something if I show you some interesting figures, which shows if I power versus delay tau, then when it is p tau constant, we are essentially running on this line, which essentially means, if you have a larger delay that is slower speed and this is your power if you have a smaller delay, obviously power is larger, if you have a larger delay or slower circuit you have a lower power.

This is essentially the figure of merit because once now whatever you do, you can move on this line only and therefore larger the power smaller the delays, larger the delays smaller the power. However, the efforts should be may be some other technology may or somehow method we may have larger  $p \tau$ , you may have even larger  $p \tau$ , or you may have technology in which you have reducing  $p \tau$ .

And if that occurs one can say for a smaller delay, you will have relatively compared to this smaller delay itself, you have a smaller power. So the effort in low power circuit is not only to look at all other options, depending on the topology, switching transient something, we also look into different technology nodes, different technology method like using, one can say using (()) (39:30) hence you can use it using (()) (39:32).

Or you can use other semiconductor technologies to reduce this  $p \tau$  lines from this. Let us say if this was your CMOS, standard silicon, I want to go below, so I must do something in which PDP remains, which is essentially joules this energy PDP is essentially. So this is the constant energy graph shown to you here per switching event and therefore the effort in the technology are effort in the all low power circuit designees to reduce if possible in a  $p \tau$ .

But in a standard CMOS circuit  $p \tau$  will remain constant and therefore the worry start, that if then if I want to reduce power and do not want to give speed, that is I want to keep TDP constant, what are the methods or what can I do that I can fool the circuit and there it can say I have lower power, lower energy consumed.

**(Refer Slide Time: 40:43)**

### Dynamic Power Consumption (related to Battery Power)

- Average power consumption by a node cycling at each period  $T$ :  
(each period has a  $0 \rightarrow 1$  or a  $1 \rightarrow 0$  transition)
 
$$P_{switching\_battery} = \frac{E_{cycle}}{T} = C_0 V_{DD}^2 f_{CLK}$$
- Average power consumed by a node with partial activity  
(only a fraction  $\alpha$  of the periods has a transition)
 
$$P_{switching\_battery} = \alpha C_0 V_{DD}^2 f_{CLK}$$



Activate Windows  
Go to PC settings to activate Windows

Now let us look at the dynamic power once again which is related to battery, we can see that average power construction by node cycling at each period  $T$  that is each period we call it, first transition 0 to 1 and the next is from 1 to 0, that means you are going from 0 to 1 and returning to 0 is called one transition or one period of a transition. So the switching which is we have just calculated switching cycle energy which is called  $E_{\text{cycle}}$  which is  $C_0 V_{DD}^2$  square.

And if you have  $T$  cycles in which the average has to be figured out, we divided by  $T$  and  $1$  upon  $T$  is nothing but the clock frequent  $T$  is the period means  $1$  upon  $T$  is the clock frequency which is changing the data over the period of the pulses coming in, so we say the switching battery power is  $C_0 V_{DD}^2$  into  $f_{CLK}$ , but this is essentially coming from the battery.

Now average power consumed by a node with partial activity, I already talked about this term  $\alpha$  which is called the activity coefficient. If  $\alpha$  is not one, that is every clock cycle there is no transition, that means, one cannot say that every data which is let us say, I have 1110, so the first input is 0, it gives some power, the next transition is, next input is 1, so there will be a transition from 1 to 0 at the output and 0 to 1 at the input.

However, next two bits, 1 and 1 will not change input or not change the output, and therefore there will not be any charging or discharging of capacitors and therefore there will be activity will be one out of the 4 possibilities and therefore  $\alpha$  will be very small, maybe one fourth, sometimes one third, sometimes half depends on the how much activity it happens in a given cycle.

Now this essentially  $\alpha$  is the only thing which probably we should look very carefully in our power design, low power design because this we cannot weigh too much, this of course is essentially decided by  $W$  by  $L$  which we have kept for transistors for noise margin consideration, for speed consideration. So if  $f_{CLK}$  somehow connect, is already fixed, then the only way switching energy can be minimized or power supply has to give lower energies to make  $\alpha$  smaller.

**(Refer Slide Time: 42:58)**

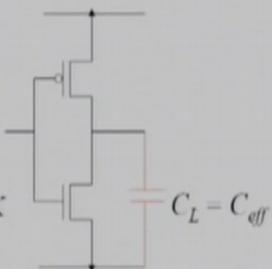
## Dynamic Power Consumption

- Define effective capacitance  $C_{eff}$ :

$$C_{eff} = \alpha C_0$$

$$P_{switching\ inverter} = \frac{1}{2} C_{eff} V_{DD}^2 f_{CLK}$$

- To minimize switching power
  - Reduce  $V_{DD}$
  - Reduce  $C_{eff}$



So we started with this slide. We say therefore the capacitance which this output node has received for an inverter is alpha times  $C_0$  and the switching power then can be half instead of  $C_0$  we write  $C_{eff}$ ,  $V_{DD} f_{CLK}$  square. So we come back now and say to minimize this switching power or dynamic power, all that we can do is to make this  $V_{DD}$  across power supply which is charging this capacitor to full  $V_{DD}$ .

If that is reduced, square times the power will go down under root of that, or since the clock we want to increase, we cannot do much on this. Only way it can reduce switching power is reducing the  $C_{eff}$  and reducing  $C_{eff}$  essentially because this is also constant, so only thing we reduce is the activity coefficient which is alpha.

**(Refer Slide Time: 43:50)**

## Factors Influencing $C_{eff}$

- Circuit function
- Circuit technology
- Input probabilities
- Circuit topology

So there are four factors which influences alpha, I was just talking of one of them, one of course is called circuit function, the other is circuit technology, the third is input probabilities and the fourth is the topology of the circuit you lay. So depending on the four such functions C effective can be different and if the C effective is different, then the power dissipation will be different.

**(Refer Slide Time: 44:20)**

**Some Basic Definitions**

- **Signal probability** of a signal  $g(t)$  is given by
 
$$P(g) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} g(t) dt$$
- **Signal activity** of a logic signal  $g(t)$  is given by
 
$$A(g) = \lim_{T \rightarrow \infty} \frac{n_g(T)}{T}$$

where  $n_g(t)$  is the number of transitions of  $g(t)$  in the time interval between  $-T/2$  and  $T/2$ .

Before we continue with this activity part, we like to first define few basic terms. Two terms of interest for us, in a case of data flowing from input to output which we call signal, if we call a term signal probability of a signal  $G$ , is given by  $P_G$  which is probability term of  $G$  of the signal, can be if  $T$  is the period then limit of  $T$  tends to infinity,  $1$  upon  $T - T$  by  $2$  to  $+ T$  by  $2$   $g$   $dt$ . Now this essentially means  $g$  essentially is the signal, it is  $0$   $1$   $1$   $0$  whatever it is.

So if you average it in any one time period that is called the probability of getting  $0$  or probability of getting  $1$ . So this is  $P_G$  which is most important because signal can be either  $0$  or one, so one can get a probability of getting output one or output  $0$  by using this kind of simple calculations or simple averaging.

The another term which we use often in the power dissipation valuations for  $C$  effective or alpha is signal activity of a logic signal  $g$  is given by  $A_g$  which is limit of  $T$  tends to infinity,  $1$  upon  $T$  again because averaging you are doing,  $n_g$  of  $T$ , now  $n_g$  is essentially we have called activity is the word we have said,  $n_g$  is number of transitions of  $g$  in the time interval between  $- T$  by  $2$  to  $+ T$  by  $2$ .

How many times input changes or output changes is called its activity. So we have a signal activity of a logic signal  $g_t$  given by  $A_g$  limit of  $T$  tends to infinity and  $g_t$  by  $t$  whereas we have signal probability of a signal  $g_t$  is given by limit  $T$  tends to infinity even upon  $-T$  by  $g_t dt$ . So why we are looking because we want to calculate the probability of any gate for the transitions.

And if that we evaluate then we know how much is alpha and if we know our alpha we know our  $C$  effective. So here is before we go for other things. Quickly few of the things may be of interest to you.

**(Refer Slide Time: 46:39)**

**Factors Influencing  $C_{eff}$ : Circuit Function**

- Assume that there are  $M$  mutually independent signals  $g_1, g_2, \dots, g_M$  each having a signal probability  $P_i$  and a signal activity  $A_i$ , for  $i \leq n$ .
- For static CMOS, the signal probability at the output of a gate is determined according to the probability of 1s (or 0s) in the logic description of the gate

Diagram illustrating signal probabilities for basic logic gates:

- Inverter: Input  $P_1$ , Output  $1 - P_1$
- AND gate: Inputs  $P_1, P_2$ , Output  $P_1 P_2$
- OR gate: Inputs  $P_1, P_2$ , Output  $1 - (1 - P_1)(1 - P_2)$

Assume that there are  $n$  mutually independent signals, that is  $g_1, g_2, \dots, g_n$  are, that means if  $g_1$  changes,  $g_2$  up to  $g_n$  need not change or vice versa, so each is independent signal which is normally will appear, data may appear independently from variety of inputs. Now, each have a signal probability  $P_i$ , that  $p_1, p_2, p_3, \dots, p_n$  and a signal activity  $A_i$  which is  $i$  has to be less than  $n$  where  $n$  is the number of input gates, input of that.

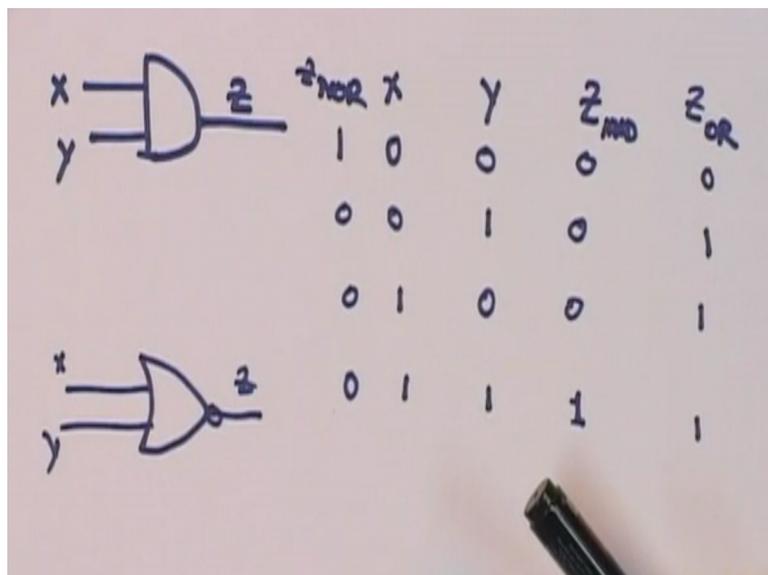
Now we define further for the static CMOS the first, because there are different variety of CMOS possibility starting dynamic, among dynamic also we have variety of choices, domino CMOS, modified dominos, zipper and a variety of them. So we first look for static which is most basic CMOS circuits and which gives you basically the idea how to reduce or how to understand the CMOS circuit behavior.

So for a static CMOS, the signal probability at the output of a gate is determined according to probability of ones or 0s in the logic description only. For example, ELN inverter, if you have 1 here, you are going to get 0. But since this is a binary number 0 is nothing but  $1 - 1$  and therefore we say 0, so we say probability of occurring at input 1 at inverter then the probability of getting a 0 at the output is nothing.

But  $1 -$  probability of occurring 1 at this input. Please remember 0 is complementary of 1 and therefore in the binary system, therefore if this is the probability of occurring one,  $1 - p_1$  is the probability of occurring 0 at the output. So inverter is essentially acting like an inverter. If there is a probability of 1 here, to occur then there will be  $1 - p_1$  probability is to occur 0 at the output due to inverting action.

Look at the gates, you have a simple two input line, and let us say  $p_1$  is the probability of having 1 at the input or  $p_2$  which is probability of having 1 at the input, then the probability that 1 will be occurring will be  $p_1$  multiplied by  $p_2$ . Now this can be understood easily.

**(Refer Slide Time: 48:57)**



You can see from here, in an AND gate, if you are two inputs, you have an AND gate, x and y and you have z. So if I write it to table, x, y and z, possible combination is a 2-bit number. So 0,1,1,0,1,1 and it is an AND function, so this is 0, this is 0, this is 1. So you can see from here, AND gate has only one probability, only 1 out of 4 possibilities of getting 1 only when inputs x and y are 1.

Whereas for any other inputs combination 00, 01 or 10,  $z$  is the output going 0. So if we have a probability of  $p_1$  getting one, is this and  $p_2$  getting 1 then the  $p_1, p_2$  is 1. If anyone of the probability is for  $p_1$  0, then the output will be 0. So essentially we know this is like a Boolean algebra, probabilities multiplies, so for an AND gate this is intersection of  $p_1$  and  $p_2$ , therefore we say if  $p_1$  is the probability of first input having 1.

If  $p_2$  is the probability of getting second input 1, then the probability of occurrence of 1 at the output is  $p_1$  multiplied by  $p_2$ , which is the product of two probabilities and which is visible from here that if you have only 1,1,1's if any one of them is 0, the output is going to be 0 which is what AND gate function is. By similar logic one can see if we have a OR gate to input OR gate again  $x, y, z$ .

And I will make again table, and we say  $Z$  OR, and now we know  $x$  and  $y$  for an OR gate, 0,0 is 0, 0,1 is 1,1 is 1. In an OR gate occurrence of 0 is very little where as occurrence of 1 is more likely to have except only once when  $x$  at 0 and  $y$  0 the output goes to 0. So if you look at the probability,  $p_1$  and  $p_2$  are the probability of occurrence of 1 at inputs, then the occurrence of output being 1 is probability  $1 - (1 - p_1)(1 - p_2)$ , and  $1 - p_1$  is a 0 and  $1 - p_2$  is another 0.

So whenever there is one of them is one, the other turn will essentially,  $1 - p_1$ , let us say  $p_1$  is 1 then that becomes 0,  $p_2$  is 1 that becomes so, so  $p_1$  output is 1. For all other cases one can see the output may become 0 which essentially means the probability of occurrence of 1 is  $1 - (1 - p_1)(1 - p_2)$ ,  $1 - p_1$  multiplied by  $1 - p_2$  which is occurrence of 0. So you subtract the occurrence of 0s out from the occurrence of 1 probability wise.

So it is  $1 - (1 - p_1)(1 - p_2)$ , which are the probability of getting 0s due to  $p_1$  and  $p_2$ , you subtract that probability, the remainder is getting a 1 at the output. Now this is very important because now we are trying to see that the probability of which kind of inputs you have will decide what is the probability of getting 1 or 0. And if there are no transition therefore required, then in that case there will not be any power consumptions. So let us look at some numbers, some theory.

**(Refer Slide Time: 52:43)**

## Factors Influencing $C_{eff}$ : Circuit Function (Static CMOS)

- Transistors connected to the same input are turning on and off simultaneously when the input changes
- $C_i$  of a static CMOS gate is charged to  $V_{DD}$  any time a  $0 \rightarrow 1$  transition at the output node is required.
- $C_i$  of a static CMOS gate is discharged to ground any time a  $1 \rightarrow 0$  transition at the output node is required.

**NOR Gate**

Again if you have a NOR gate shown here, then we see, this is  $x_1$ ,  $x_2$  for a p channel inputs and  $x_1$ ,  $x_2$  in an OR mode for the n channel inputs and we call this node 1, this node 2, this node 3 and we are taking an output at 3 and this is the 0th node which is ground. Now what is a NOR function, as long as any one of this input is 1 the output is going to be 0 only when both of them is 0, the output at that time both p channel will be on and therefore output may go to 1.

If any one of them is one, the VDD is taken away. If both of them are 0 or off, only then 0 is taken away. So it is more likely that if either  $x_1$  is 1 or  $x_2$  is 1 both are conducting. If  $x_1$  is 0 but  $x_2$  is one, this will conduct. If  $x_2$  is 0,  $x_1$  is one, so this will conduct, but in either case, either this conducts or this conducts. So power supply is removed for the three cases from the output node whereas in the case when  $x_1$  is 0 and  $x_2$  is 0.

There is no path from output to the ground, fortunately because the complementary action both p channel receives input 0, 0 which makes this transistor turn on and VDD is then transferred to it, 3 or 1 is transferred to. So we now see, the transistor connected to the same input are turning on and off simultaneously when the input changes. Please remember, transistor connected to the same input are turning on and off simultaneously when the input changes.

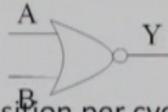
Thus, capacitance load of static CMOS gate is charged to VDD at any time when the 0 to 1 transition occurs. When 0 to 1 transition occurs, the capacitor is pre charged to this. CL of a

static CMOS gate is discharged to a ground by any time 1 to 0 transition, occurs if this was 1 and we want to go to 0, either through this or both together I can bring this to turn to 0.

Now you can see the effective capacitance at x3 in the different kind of inputs will have different kinds of are available to it, essentially means the time constraint will be different for each cases and therefore the currents will be different and if the currents are different then the dynamic power will be different.

**(Refer Slide Time: 55:15)**

### Factors Influencing $C_{eff}$ : Circuit Function (Static CMOS)

- Two-input NOR gate 
- Assume only one input transition per cycle is allowed
- Assume inputs are equiprobable:  $p_A=p_B=1/2$ .
- The probability for the output to be 1 is  
 $p_Y=(1-p_A)(1-p_B)=1/4$  
- The probability for the output to be 0 is  
 $p_{Y'}=1-p_Y=3/4$

Take another example from here, we see you have a two input NOR gate which is shown here which has two inputs A and B and an output Y. Assume only one input transition per cycle is allowed, that is A may go to 1 and that time B may not in that transition, so allow only one input transition per cycle is allowed. Assume inputs are equi probable. That means there is a possibility of change of inputs is equal, that is getting 1 on A, 1 of B is equal.

Because they are same inputs, they may not be but in this case we assume they are same, so we say each has a probability of occurrence of half. The probability is half because other time it may be 0. Either it is 1 or otherwise it is 0 and therefore the probability is half for both PA and PB, the probability for the output to be 1 which we call PY is since it is going to be NOR function.

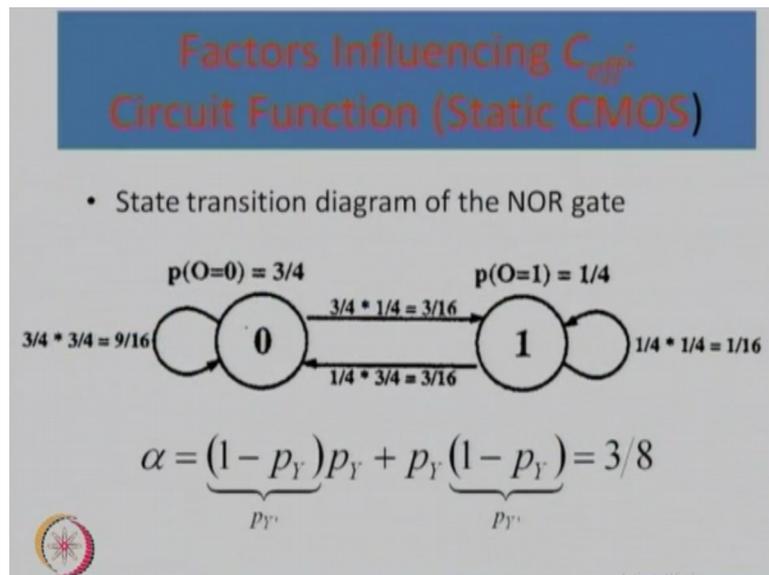
So you must get probability of occurring 1 is,  $1 - P_A$  is for the 0,  $1 - P_B$  is for the 0, which 0 probability we find out, for that this is  $1 - P_A$  is half,  $1 - P_B$  is half and therefore output becomes, probability is 1 by 4 at Y. However, if the probability of finding a 0 at the output

you are looking for then just subtract  $1 - P_Y$  from this we calculate it, so remainder probability of getting 0 is 3 by 4.

You can get from here, for the same circuit, if I put ZNOR for this, now one can see this is what is the ZNOR. Since this ZNOR, you can see out of 4 you have probability of getting three 0's. So you have a probability of getting 0 is 3 by 4, whereas probability of getting 1 is only 1 out of 4. So it is 1 by 4. So why we use, we can always see it from the two table, however since the functions may not be 2 input or 3 input, maybe larger inputs.

Evaluation of such probabilities is best left to a simple formulation rather than actually looking at the truth table and then finding the value, is that clear? For ZNOR occurrence of 1 out of the four possible output is only 1 by 4, whereas occurrence of 0 is 3 times that, so 3 by 4. So therefore one knows that a probability of output to be 0 is 3 by 4, probability of occurring a 1 is 1 by 4 for a NOR gate.

**(Refer Slide Time: 58:03)**



We actually now look into factors influencing C effective, continuing with that and we still believe that we are working on static CMOS. Here is what the same NOR gate representations can be shown on a finance state machine kind of transition table or transition diagrams. Let us say there are two states in the system, you are in a state 0 or you are state at this state at 1 . And O is the output.

So if you see, you are in a state 0, then you are making a transition to state 1 . Please remember 0 to 0, 0 occurrence has a probability of 3 by 4, 1 occurrence has a probability of 1

by 4. So we say when you are making a transition between 0 to one, the net probability is 3 by 4 multiplied by 1 by 4 which is 360. If you are going a transition from 1 to 0 which is again same, 1 by 4 multiplied by 3 by 4 is 3 by 16.

However, if you are making a transition from 0 to 0, since 0 has a probability of three fourth, then the probability of going from 0 to 0 is 3 by 4 multiplied by 3 by 4 which is 9 by 16, by same logic if probability is if you are going from transition of 1 to 1, then the probability of occurrence is 1 by 4 into 1 by 4 is 1 by 16. So obviously you can see getting 0s are much larger probability, 9 by 16 compared to 1 by 16 for one to one transitions.

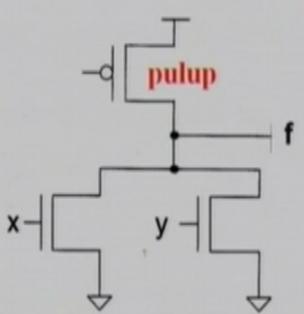
So if we want to define our activity coefficient, one can say that the occurrence PY dash which is  $1 - PY$  times  $PY$ , +  $PY$  times  $1 - PY$  and if I substitute from here,  $3$  by  $16$  +  $3$  by  $16$  then it is alpha is  $3$  by  $8$ . So I can evaluate the activity coefficient of a NOR gate which as I said by the truth table has the largest probability of getting 0 whereas smaller probability of getting one, the activity coefficient is  $3$  by  $8$ th.

Please remember this  $3$  by  $8$ th will now get multiplied to  $C$  effective and all our effort is to see for different kinds of gates, we may evaluate alpha and if which ever gives smaller alpha then one can see from here, smaller is the alpha, smaller is the  $C$  effective and smaller is the switching power. So effort in design, you can now understand is to see to it that the transitions are such that, or the gates you use are such that their activity coefficient for the same data inputs is smaller and that is one way of reducing the power dissipation.

**(Refer Slide Time: 01:00:58)**

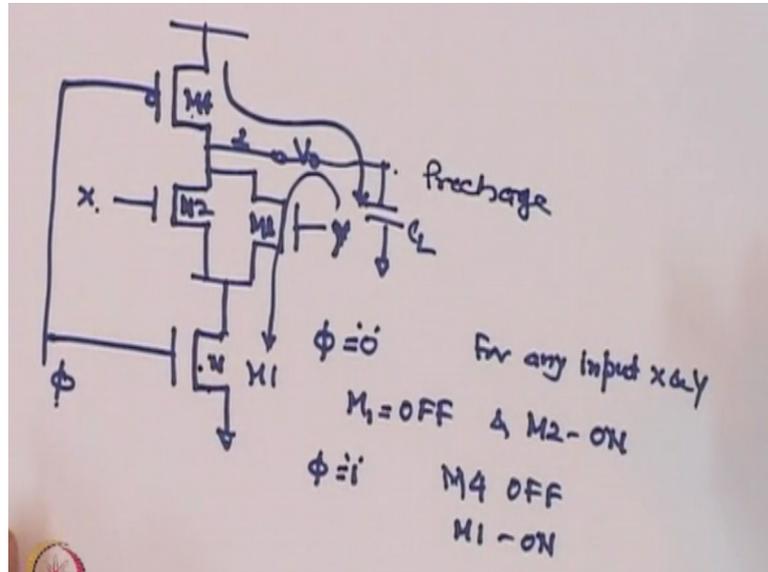
### Factors Influencing $C_{eff}$ : Circuit Function (Dynamic CMOS)

- At each cycle, PUP is precharged to  $V_{DD}$ .
- $C_L$  is precharged to  $V_{DD}$  at each clock cycle
- It is discharged to ground any time a  $1 \rightarrow 0$  transition at the output node is required.



Now take a dynamic circuit. This we have talked about so far only static but when can see from here a dynamic circuit, a slightly better dynamic circuit has been shown here, typical CMOS dynamic circuit looks something like this.

**(Refer Slide Time: 01:01:12)**



You have p channel device, you have for a NOR gate you have two n channel devices, two inputs x and y and then you have another CMOS device and NMOS device and this is your PMOS which is dynamically controlling through a clock which is phi, this is your power supply and this is your Vout, of course you have a capacitance here. Now one can see from here, the z output here is decided by dynamic circuit in the same sense.

Because when phi is equal to 0, this n channel is off, okay? When phi is equal to 0, we call it say M1, M2, M3, M4. For any input x and y, if phi is 0, M1 is off and M2 is on. Since this is off and this is on and this may conduct or may not conduct depending on x and y value, but the ground is removed because phi is 0 and if that occurs, there is this current may start from power supply to M4 and will charge the load capacitance.

So you have a charging transient when phi is going from 0 to 1. So it will go towards VDD. When phi goes to one and here x and y can now change or earlier whatever they can now change and if that occurs between phi 0 and phi one, then now M4 is off, 1 means logical ones, okay? M4 is off, M1 is on now, fully on. So this depending on the logic requirement, whether you want a 0 or one.

If let us say you want 0, that means x must be 1 or y must be 1 or both maybe one, only then one of the two transistors or both of them will be on and the capacitance will discharge through this part. But if say x is equal to 0 and y is equal to 0, M2, M3 will be switched off anyway and for whatever this charge it was retained in the charge which was VDD due to the VDD, therefore 1 will be retained.

So now you can see in a clock cycle phi is equal to 0, this is dynamic in that sense, the power is decided by, it is not just decided by this logic, but also decided but the timing given to phi and because this is dynamically at no time these and these two are on, that is M4 and M1 are on, clearly one sees that the power dissipation will be smaller, at least transition, there will not be any switching because at no time, all four will be actually acting.

And therefore this chain will always be broken and therefore there will not be any what we call as this short circuit power. Now this is called pre charge mode verify is 0, M2 is on, we precharge to VDD and when depending on x and y. Now this the circuit shown here in my figure you can see I have removed the lower n channel. The reason why we remove n channel, anyway n channel was providing you a ground path.

So we kept them ground directly to the AND and NOR, the only problem a little more power it may consume once a while because if either of the input x and y is 1 and even if your phi is 0, part of the power will go to the ground through x or y inputs or those transistors because any one of them is 1 logically. So then the important thing there is that whenever such circuits are used.

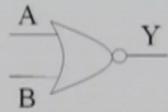
The inputs are always initially 0 and they are changed when phi goes to 1 for the pull up. If that occurs the power pre charge occurs when phi is 0 and when phi is equal to one, the pulup turns switches off and the logic is as per x and y. Now this dynamic CMOS has lower number of transistors, sometimes little more short circuit power it may create but it has overall low power compared to many other dynamic gates.

So this circuit which is a NOR gate again, which as I say the lower transistor has been taken off, can also give you comparatively lower C effective because now the transitions will be smaller and for a NOR gate and if that occurs the net C effective with a lower per cycle transition can lead to a lower C effective and therefore low power dynamic power. So this is

very interesting that you use a dynamic power, a dynamic circuit then you have much lower compared to lower power compared to static CMOS.

**(Refer Slide Time: 01:06:54)**

### Factors Influencing $C_{eff}$ : Circuit Function (Dynamic CMOS)

- Two-input NOR gate 
- Assume only one input transition per cycle is allowed
- Assume inputs are equiprobable:  $p_A = p_B = 1/2$ .
- The probability for the output to be discharged is  $p_Y = 3/4$
- The probability of  $C_L$  to be re-charged at the next cycle is  $p_Y'$ .

If you are using a dynamic CMOS and you have again the same circuit as I have shown here and we still assume that they have a occurring 1 at here is probability of half and we also now in the case of dynamics there is only 1 input transition per cycle is allowed. You cannot change all of them, only one at a time. The probability of output to be discharged is 3 by 4 because one of the input maybe 1 and therefore it will create a path to create a three fourth.

If both of them inputs are 0 only then the output is high remains. The probability of CL to be recharge in the next cycle is  $p_{\gamma}'$  which is nothing but  $1 - P_{\gamma}$ ,  $P_Y$ . So essentially it is one fourth. So in nutshell what we can say about the dynamic circuit versus static circuit, following things can be thought of.

**(Refer Slide Time: 01:07:40)**

### Factors Influencing $C_{eff}$ : Circuit Function (Dynamic vs. Static)

- $\alpha_{dynamic\ CMOS} \geq \alpha_{static\ CMOS}$ :
- $C_{eff\ (dynamic\ CMOS)} \leq C_{eff\ (static\ CMOS)}$
- Power due to glitching is much smaller in dynamic CMOS than it is in static CMOS.
- In static CMOS, the transition probability depends on both input probabilities and previous state.
- In dynamic CMOS, the transition probability depends on solely input probabilities.
- In static CMOS, the gate output does not switch if the inputs do not change between subsequent cycles.
- In dynamic CMOS, the gate output may switch even if the inputs do not change between subsequent cycles.

We say alpha dynamic CMOS, alpha means activity coefficient for dynamic CMOS may look to be little larger compared to static CMOS, however the C effective dynamic CMOS even if because the net capacitance is only smaller in this case, so C effective for static CMOS is always larger than C effective for dynamic CMOS. So what is our important is that C effective should be reduced.

So in a dynamic CMOS circuits, since we are reducing C effective, overall we figure it out that it can become little low power compared to static CMOS. Now power due to glitching is much smaller in dynamics. The glitch does not occur because all transitions are not allowed during the fire, when the phi is constant. The phi is 0 or phi is 1 only then the inputs are allowed to change.

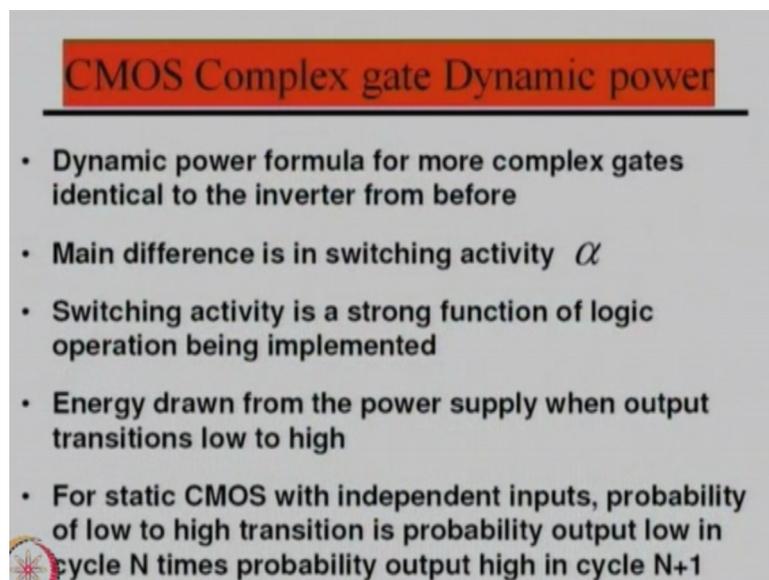
And because this is a constraint at no time VDD to ground path can be created, so there is no glitch is possible in a dynamic CMOS which in a static CMOS it can occur because as soon as any input changes, both n channel and channel n complement p channel will change its status and because of that there is a possibility of glitch. In static CMOS the transition probability depends on both input probabilities.

And the previous state, we have just calculated that. In the case of dynamic CMOS, the transition probability is essentially only depends on the input probability and nothing to do with the output probability because you are already in the precharge mode, either to remain charged or to discharge is only one possibility and therefore one can say that in a dynamic

CMOS transition probability depends only on the inputs and not at the output. In static CMOS the gate output does not switch if the inputs do not change.

That is most important, okay, between the cycles. In dynamic CMOS the gate output may still switch if the inputs do not change between subsequent cycles because 5, 5, 0 that is p channel and channel dynamic part may still switch on and off and the output may still go to 0 to 1 or 1 to 0 in either case. So there is some disadvantage in dynamic CMOS that inputs change, even if inputs do not change, output can change whereas in the static unless the input changes output cannot change.

**(Refer Slide Time: 01:10:16)**



**CMOS Complex gate Dynamic power**

- Dynamic power formula for more complex gates identical to the inverter from before
- Main difference is in switching activity  $\alpha$
- Switching activity is a strong function of logic operation being implemented
- Energy drawn from the power supply when output transitions low to high
- For static CMOS with independent inputs, probability of low to high transition is probability output low in cycle N times probability output high in cycle N+1

By similar logic, one can use a complex gate, not simply NAND, NOR or anything, any of the complex circuits using CMOS and we can always by logical effort can see its equivalent inverter action and then for can calculate the alpha. The main difference is the switching activity for any complex gate. So we must evaluate alpha for such gates.

Switching activity is strong function of logic, so which logic you are implementing that decides the switching activity and energy drawn from the power supply when output transitions is low to high, that is a charging we can see, for static CMOS please remember power supply only gives in the half cycle when it is charging the capacity. So energy is only drawn when you are pre-charged in your pre-charged system.

And during the evaluation phase there is no power supply requirements and therefore you are reducing the power supplied by the power supply in the case of dynamic CMOS. For static

CMOS with independent inputs, probability of low to high transition is probability of output low in N times probability output high in the N + one time. This we shall see by example and you will understand what I meant.

**(Refer Slide Time: 01:11:30)**

**CMOS Complex Gate Activity Factor**

**Fundamental Activity Factor Equation:**

$$\alpha_{0 \rightarrow 1} = p_0 p_1 = p_0 (1 - p_0)$$

- Assuming inputs are independent, uniformly distributed, then any N input static gate has transition probability:

$$\alpha_{0 \rightarrow 1} = \frac{N_0}{2^N} \frac{N_1}{2^N} = \frac{N_0 (2^N - N_0)}{2^{2N}}$$

- $N_0$  is the number of 0s in truth table output column
- $N_1$  is the number of 1s in truth table output

Now here is a, for any complex gate activity, we say alpha is activity coefficient for CMOS complex gate alpha going from 0 to 1. If you have probability of getting 0 is P0 and probability of getting 1 is P1, and we know P1 is nothing but 1 - P0, so we say alpha 0 to 1 is P0 1 - P0. Now again same assumptions we make, assuming inputs are independent and they can be separately controlled.

And they are uniformly distributed in and input static gate has transition probability from 0 to one where if N0 is the number of 0es and N1 was the number of ones in the truth table column, output column that is Z or Y whatever we say, then we can say from here, N0 by 2 to the power N multiplied by, out of 2 to the power N, let us say you have a two bit number, so N is 2. So you have 2 to the power 2 means 4.

If number of 0s are, let us say your 2 bit possibility is you have 0,0,0; 1,1,0; 1, 1, so one can see N0 divided by, how many 0s you have at this columns divided by the 4 because as I said last time in a NOR gate we have 3 0es and one 1 or vice versa OR gate. So in which case out of 4 possibilities, 1 0 or three times 1 is possible. So the product of getting 0 or 1 out of 2 to the power N possible combinations is for alpha.

Since N1 is nothing but number of ones in truth table which must be total possible combination of bits minus the N0 2 to the power of this. So I now can evaluate can any input, let us say N is 2, 2-bit number, the 2 to the power 2 is 4, one can see from here, so it is 2 multiplied by 2 to the power 4 that is 16, let us say number of 0s are 2-3, then this is 2 to the power 2 is 4, 4 - 3 is 1.

So it is 3 by 16 is the probability of alpha going to 0 to one, by similar logic for a NAND gate it will be N0s would be different. So whatever number of inputs you have, whatever complex gate you create, alpha 0 to 1 and similarly alpha 1 to 0 can be attained in this simple formulations. Please remember N0 is number of 0s in the truth table output column that is why an N1 is number of one's into truth table at the output.

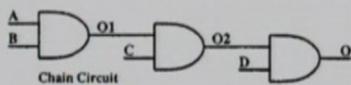
Since I can therefore fundamental activity factor alpha for 0 to 1 or 1 to 0 can be found per cycle. I can find out the c effective and if I know my c effective I can get mine. We may actually stop here for today but we may just give you the last slide which I will repeat next time. First time we looked into whether to use static CMOS or to use dynamic CMOS, to increase the C effective, that is reduce the C effective or impact we looked into.

Now next thing which we will like to look into is the topology of the circuit has anything to do with the C effective.

**(Refer Slide Time: 01:14:59)**

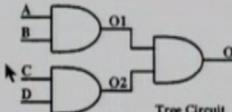
## Factors Influencing $C_{eff}$ : Circuit Topology

- Circuit topology may have high impact on  $C_{eff}$
- Example: Chain and Tree implementation of a four input NAND gate
- Assume static CMOS
- Assume all inputs are equiprobable.



Chain Circuit

$\alpha_{O1} = 3/8$   
 $\alpha_{O2} = 7/32$   
 $\alpha_O = 15/128$



Tree Circuit

$\alpha_{O1} = 3/8$   
 $\alpha_{O2} = 3/8$   
 $\alpha_O = 15/128$

To some extent we have worked out this kind of chain and tree circuit in our logical effort part. Now we will like to see this and this may give you same logic, however we will like to

see what is the activity coefficient alpha for this and what is the activity coefficient for this. If this gives you lower alpha or this gives you lower alpha then we can use one of the chains or tree circuit to implement your logic.

Please remember logic remains same from input to output in either case, either you can represent in this tree form or you can have a chain of circuit as shown here and the output can still be same function as we are looking into. Now this is a 4 input 9 gates you can have, you have 2 input, 2 input, or you have 2 input, 2 input, 2 input. So therefore inputs A,B,C,D. As I say AB here or CD here or AB first then output with C and then output of this with D and then final output and we right now take a case when assume all static CMOS.

Assume all inputs are equi probables and can evaluate the alpha for these circuits. Even if alpha is we are just, I will do it again, they seem to be same for either of the case, but one interesting feature of these two circuits probably if you have done logical effort properly we have seen that if you have a chain of circuit, the delays are larger whereas because this has to wait to evaluate C into this there will be delay from this whereas in this case this is equal delays.

This and these two will come together and therefore the delay is minimum. This is the 3-unit delay system. This is 2-unit delay system, so in E1 if their power is essentially similar, C effective is similar at least the speed wise this circuit will be faster compared to this. So depending on the power and speed requirements we not only look into the which kind of gate, static or dynamic or complex dynamics we should use.

Please remember dynamics have many possibilities as I said, one is of course these simple dynamics then we call it, domino, then we called modified domino, we have a (( )) (01:17:19), we have zipper, any possibility of dynamics CMOS can be tried and we will look into at least some of them later, other possibilities of different way of doing the same logic. So with this we will conclude for the day and we will come back next time.

And we will start again influence of C effective on the circuit topology and influencing C effective which coming from circuit topology and will see whether which one of this circuit should be chosen so that of course there are issues which this slide is showing which I will

discuss is called as skew problems. So if the data is one particular and you may land into skew though you may get better result otherwise.

With this kind of circuit, we will look into the power again. And therefore we stop here today and we have learned so far in this is that there are three kinds of power or rather four sometimes but at least 3, one is the dynamic power, the second is the short circuit power and third is the leakage power. We are still working right now on the dynamic power. And to some extent into automatically we are also looking into switching power, I mean short circuit power.

We will look into more details about the leakage power separately. With this thank you for the day and we will meet next time. Bye.