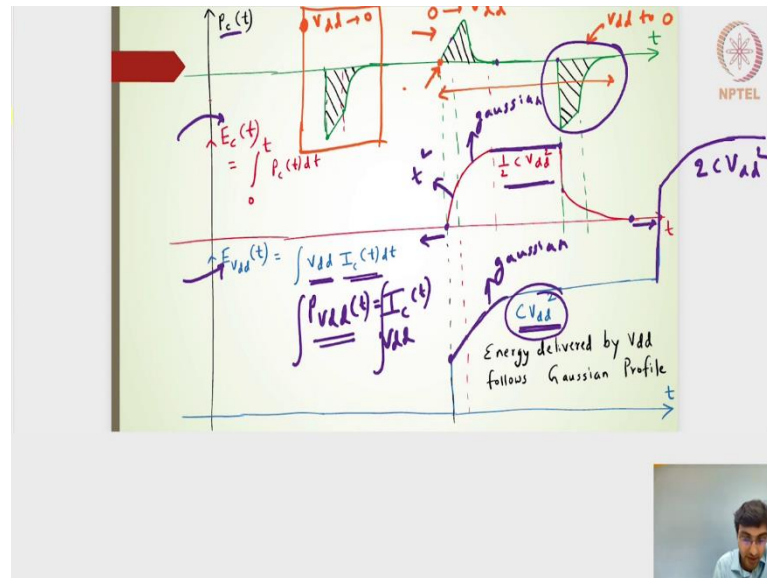


**Design and Analysis of VLSI Subsystems**  
**Dr. Madhav Rao**  
**Department of Electronics and Communication Engineering**  
**International Institute of Information Technology, Bangalore**

**Lecture - 62**  
**Switching Power and Energy Estimation**

(Refer Slide Time: 00:15)



Hello students, welcome back to this lecture. What we have been looking in this particular lecture is trying to understand the power across the time domain for an inverter which is actually switching at the output side it is actually you know going it is rising to 1 and then coming back to 0. In this particular case what we are saying is the power which is kind of consumed by the capacitor while the capacitor is getting charged from 0 to  $V_{dd}$  that means that the output voltage here is rising it is going from 0 to  $V_{dd}$ .

In this particular case the power it is kind of the opposite of what it is consuming. The power it is kind of dissipating through the NMOS transistor, it is this particular thing we can say that the output node voltage while the capacitor is discharging from  $V_{dd}$  to 0. Here this particular profile is from 0 to  $V_{dd}$  and I am going to ignore this particular case where it is going from  $V_{dd}$  to 0.

This particular portion I am going to remove now. I am going to be interested in starting with the capacitor getting charged. The output node voltage will go from 0 to  $V_{dd}$  and then

while it is going back from  $V_{dd}$  to 0. I have a step input where it is going from the step input side it will go from 1 to 0 and then here the step input will go from 0 to 1. In that particular sense what should be the energy that will be consumed or rather it will be stored by the capacitor.

As we know the energy term is nothing but the integral of one over time of this particular instantaneous power. If I actually look into the energy term for this particular instantaneous power it is actually nothing but the addition of all the instantaneous powers at that particular time domain.

If I want to do it is basically an integration, so starting with this particular value. The energy that has been consumed or the energy that has been stored by the capacitor, it will be nothing but the integral of this particular expression and then it will be nothing but the integral of this particular expression. Across this particular time domain I should be able to draw a particular energy profile. When it is at this particular starting point when it is at 0.

We will have the 0 value here of the energy and while it is increasing and in fact you will see that for this particular profile it is linear and then it will be the integration of some kind of an exponential profile and exponentially decreasing profile. While it is kind of linear I will have a square with respect to linear in the sense it is this particular power term is linearly having a linear relationship with respect to time domain.

In this particular case when the power was linear the energy because it is an integral of the linear relationship with respect to time domain, so this will be a function of  $t$  square. It is a square of the time domain. It is going to rapidly increase over the time over the portion where it was linear and then we have this kind of exponential profile and then exponentially decreasing profile. I will have a kind of integral of the exponential profile which will form kind of a Gaussian profile here or close to a Gaussian profile.

While it reaches to 0 here somewhere here, the 0 addition to whatever is the accumulated energy here will retain that cumulated energy. Whenever the power reaches 0 we will have a 0 the constant value here till for the energy and then it rapidly drops to here. I will have a drop here and then it will be going back to 0.

I will have this energy term. Whatever is the accumulated energy this particular power in the negative side will be able to subtract this particular accumulated energy it will be able to subtract and then make it 0. The time it goes to 0 we will have the 0 value here. What this particular profile says is while the capacitor is kind of charging I will have the accumulated energy.

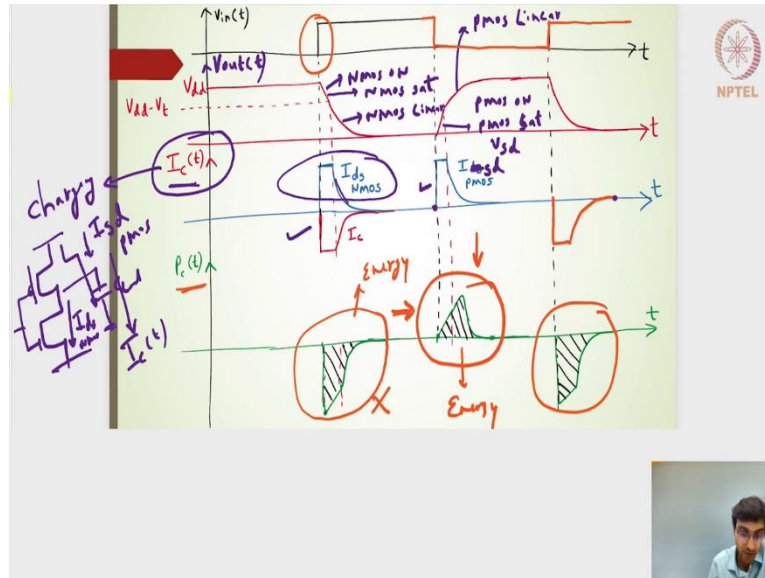
I will have the energy stored in the capacitor and while the capacitor is actually discharging the capacitor is discharging from  $V_{dd}$  to 0, the stored energy will be released through the NMOS transistor. It will be released through the NMOS transistor to the ground. This particular profile will bring back the energy that will be stored in the capacitor to 0.

Before this the energy stored by the capacitor is actually 0, after this the energy stored by the capacitor will be 0. In between while the capacitor is charging and the capacitor discharging, I will have the energy profile with respect to time domain and eventually it will be 0. For the charging I will have  $\frac{1}{2}CV_{dd}^2$  and discharging it will come back to 0.

The magnitude of this particular line while it is completely charged that means, while it is completely charged means the power is 0 here because the current will be 0 after it is completely charged, I will have this particular level as  $\frac{1}{2}CV_{dd}^2$ .

Now, I want to compare this energy of the capacitor with respect to the energy delivered by the  $V_{dd}$ . If I notice the energy delivered by the  $V_{dd}$  in terms of an expression it is nothing but an integral expression of  $V_{dd}$  which is a constant multiplied by the current of the capacitor.

(Refer Slide Time: 06:27)



If I go back to the previous slide this is my current and if I do this particular current profile was this is negative positive negative. If I consider the current of the capacitor in the charging direction I will have this positive charging one and then this negative, while it is discharging and if I consider this one and this one and multiply by a constant of  $V_{dd}$  I will have the same profile.

The power profile of the  $V_{dd}$ , the power delivered by the  $V_{dd}$  with respect to time domain will be nothing but the same profile as that of the current. If I want to do an integration of that, I need to start accumulating the points across all the time domains from at this particular point to this particular point. If I go back to the present slide.

What I have done is I have started accumulating all the points because the profile of the power of the  $V_{dd}$  will be very very similar to that of the  $IV_{dd}$  or rather the capacitor with respect to time domain. Because  $V_{dd}$  is a constant and constant multiplied by the profile  $I_c(t)$  will give me the similar profile, it will be slightly scaled because of the  $V_{dd}$  value.

If I do the integral of this particular power profile I will get the accumulated points something like this. It will start from here and then it because of the constant value of the current I will have a linear profile and then I will have an exponential profile which will be nothing but a Gaussian profile here.

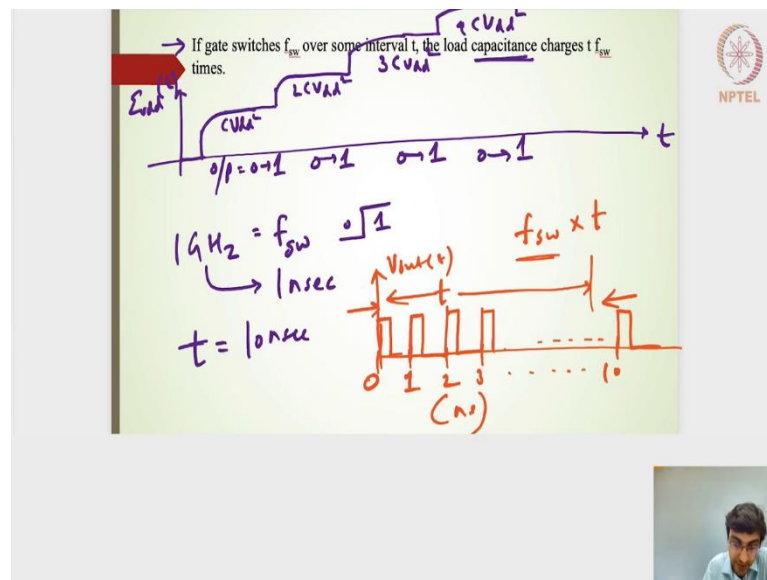
The integration of the exponential profile will give me close to the Gaussian profile. Finally, it will reach to a value which will be nothing but  $CV_{dd}^2$ . Because this is the energy

that has been delivered by the  $V_{dd}$  and this particular level is  $\frac{1}{2} CV_{dd}^2$  because this is the energy that is to be stored or consumed by the capacitor. I have written here the energy delivered by the  $V_{dd}$  follows the Gaussian profile.

What we observe here in this particular case is the energy that is kind of delivered for the  $V_{dd}$  one time and this is the energy profile. If it delivers the second time I will have something like this is the second time and similarly if I have the switching of the output voltage from 0 to 1, 0 to 1, 10 times I will have the energy delivered by the  $V_{dd}$  will follow this particular profile 10 times.

I will have  $CV_{dd}^2$  the second time at the output switches from 0 to 1 I will have  $2CV_{dd}^2$  as a level the 10th time, it is going to switch the output node voltage is going to switch from 0 to 1, I will have the energy divided by the  $V_{dd}$  as 10 times the  $CV_{dd}^2$ .

(Refer Slide Time: 09:24)



In short what I am saying is, if I have the profile of the energy with respect to time domain the energy delivered by the  $V_{dd}$ . I will have this particular profile and this particular profile and it will keep on going which will give me the value of the magnitude value  $CV_{dd}^2$ ,  $2CV_{dd}^2$ ,  $3CV_{dd}^2$ ,  $4CV_{dd}^2$  and then so on. The number of times the output switches from 0 to 1, from 0 to 1, from 0 to 1, the energy will be delivered by the  $V_{dd}$ , so that the capacitor charges to  $V_{dd}$ .

That the capacitor stores  $\frac{1}{2}CV_{dd}^2$ , but the energy that has been delivered by the  $V_{dd}$  rail will be nothing but  $1CV_{dd}^2$  and then the next time there is a switching it will be  $2CV_{dd}^2$ ,  $3CV_{dd}^2$  and  $4CV_{dd}^2$  and so on. Let us say with this particular statement if the gate switches over some time interval and then the load capacitor charges for how many times.

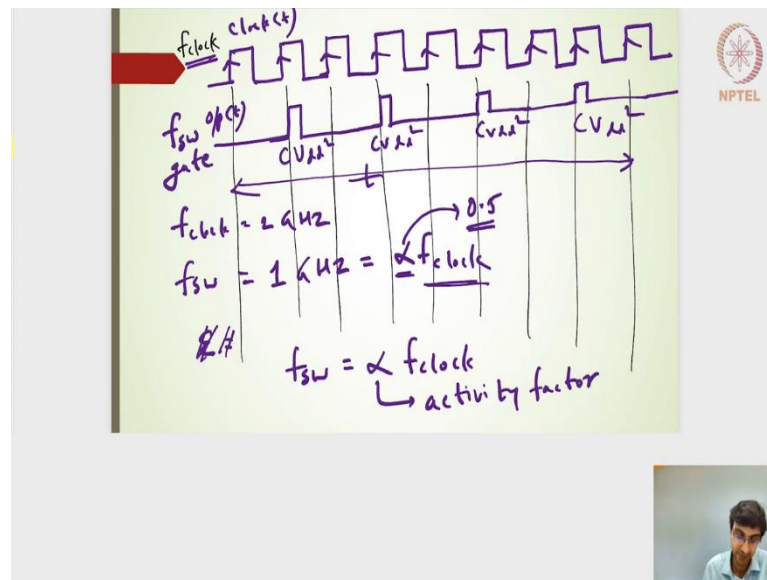
Then this is very important because just to understand what is the overall energy that has been delivered by the  $V_{dd}$ , we need to know how many times the load capacitance charges. Let us say that the output gate switches in 1 gigahertz. What I mean by 1 gigahertz is the output node switching from 0 to 1 and it happens at 1 gigahertz. In that sense if it actually switches for every 1 gigahertz means for every 1 nanoseconds there is a switching, the output switches for every 1 nanoseconds.

If my time duration if my time of interest is for a longer duration of say 10 nanoseconds. What it means is I have taken 1 from 0, 1, 2, 3 and up till 10 nanoseconds. All the scalar in nanoseconds, I am going to write it like nanoseconds the output is actually going. The output means I am going to draw the output voltage.

It switches here and then somewhere it comes back maybe and if it switches here, comes back here, switches here, comes back here and then so on it switches here and then comes back here. From 0 to 10 nanoseconds or maybe in the 0th second also I will write that it switches and then from here, till this particular portion it is going to switch how many number of times.

The number of transitions the output node does will be nothing but,  $f_{sw}$  multiplied by the time duration here at the time of interest we are having. Over a time of interest  $t$  and if I can calibrate or characterize the switching rate of that particular gate at the output side I can have  $f_{sw}$ . Hope this is clear to everyone. I am going to take one more slide here. I am going to insert one more slide new slide and then say that, let us say I have a clock.

(Refer Slide Time: 13:05)



I have a clock here of let us say 3 gigahertz or let us say 2 gigahertz. This clock is there for the overall chip design and let me take one reference the rising edge I am going to take. Then I am going to have a rough kind of a reference. So, that it becomes easier for my drawing and let us say that this is my clock frequency  $F_{\text{clock}}$ . Let us say that the output of that particular gate whatever gate we have designed it is going to switch only after every two such clock cycles.

What I meant is the switching of the gate in this case, it could be an inverter or any kind of a combinatorial circuits. Let us say that it is going to switch in this particular case it is going to switch somewhere here, so the output is going from 0 to 1. Although it is an  $f_{\text{clock}}$  it is a clock signal clock off time and here it is nothing but the output of time although I have written it as a frequency of gate. It is basically the output voltage with respect to time domain.

It is going to switch from 0 to 1 and then somewhere it is going to go back and then after this thing it will going to switch and then somewhere here. After every two clock cycles it is switching and then similarly somewhere here.

If suppose this particular clock, the  $f_{\text{clock}}$  let me say that this is 2 gigahertz. My  $f_{\text{switching}}$  is turning out to be 1/2 of that 2 gigahertz, I am going to say that this is 1 gigahertz. My overall energy which we calibrate in terms of the frequency I can actually establish an expression.

But the switching frequency can always be a function of the clock frequency. The clock is a global signal and we know the global signal or the global clock signals frequency. If I can characterize almost all the gates whatever combinatorial circuit gates we have and at the output side it is going to switch, I can have a relation with respect to that of the global clock signal. I can have a  $\alpha f_{\text{clock}}$ .

With respect to the clock signal I can have this alpha. In this case  $\alpha = 0.5$ , but generally we would not have such a fast switching or the output such a high frequency of the switching. Because if I have a very high frequency of the switching that means, that every time the energy has to be delivered by the  $V_{\text{dd}}$  at every of this switching output. Here if the switching output is there, I have to deliver  $CV_{\text{dd}}^2$  and here also  $CV_{\text{dd}}^2$ , here also  $CV_{\text{dd}}^2$  and then so on.

In this particular instance I have to actually deliver for the overall instance from here to here I have to deliver  $4CV_{\text{dd}}^2$ , but in terms of the clock signal I can write it as alpha  $f_{\text{clock}}$  and where  $\alpha=0.5$ . Using this particular expression of  $f_{\text{sw}}$ . The  $f_{\text{sw}} = \alpha f_{\text{clock}}$ , this  $\alpha$  is nothing but the activity factor. Using this particular factor we are going to establish the energy and the power expression.

(Refer Slide Time: 17:00)

If gate switches  $f_{\text{sw}}$  over some interval  $t$ , the load capacitance charges  $f_{\text{sw}}$  times.

$$P_{\text{avg}} = \frac{E_{V_{\text{dd}}}}{t} = \frac{C V_{\text{dd}}^2 (f_{\text{sw}} t)}{t} = C V_{\text{dd}}^2 f_{\text{sw}}$$

$P_{\text{avg}} = C V_{\text{dd}}^2 f_{\text{sw}}$  dynamic power

$E_{\text{avg}} = \alpha C V_{\text{dd}}^2$  1 clock cycle

$f_{\text{sw}} = \alpha f_{\text{clock}}$  activity factor

Activity factor is the probability that current node transitions from 0 to 1, because  $V_{\text{dd}}$  supplies energy only when output changes from 0 to 1.

That is what I have rewritten here if the gate switches  $f_{\text{sw}}$  times over the time interval  $t$  the load capacitance charges  $t f_{\text{sw}}$  times. The overall energy that has to be delivered by the  $V_{\text{dd}}$  will be nothing but this whole component multiplied by  $CV_{\text{dd}}^2$ . That is what I have written



here the energy that has to be delivered by the  $V_{dd}$  if it switches  $f_{sw}$  times it will be nothing but,

$$P_{avg} = \frac{E_{V_{dd}}}{t} = \frac{CV_{dd}^2(f_{sw}t)}{t} = CV_{dd}^2 f_{sw}$$

The average power delivered by the  $V_{dd}$ , the average time duration of interest is  $t$ , then I will take it as  $t$  here.

$$P_{avg} = CV_{dd}^2 f_{sw}$$

This is the average power  $V_{dd}$  is nothing but the dynamic power.

It is the dynamic power, in the sense it is nothing but the output node switching or the output node switching power also called as the switching power or the dynamic power and it is not called as a static power. Because, static power is always related to a power that has been dissipated by the circuit that has been consumed by the circuit, when the output voltage will be in a steady state.

This particular dynamic power is used for the output node while it is establishing the steady state while it is doing the transition from 0 to 1, while the capacitor is actually doing that charging from 0 to 1 that is when it extracts the power from the  $V_{dd}$ , it is called as the dynamic power and it is not the static power and we will have a look into the static power later on. This is the  $P_{avg}$   $V_{dd}$  which is nothing but the switching power also called as the dynamic power and in terms of the clock frequency we can write it as,

$$P_{avg} = P_{switching} = P_{dynamic} = \alpha f_{clock} CV_{dd}^2$$

The average energy over the clock will now become  $\alpha CV_{dd}^2$ . In the sense what I am saying is if the overall energy is nothing but  $CV_{dd}^2 f_{sw} t$ . The average energy for the particular clock we can consider it to be nothing but  $\alpha CV_{dd}^2$ .

The  $f_{sw} t$  will give me the number of switching but over a clock I can also write it as  $\alpha CV_{dd}^2$ . This will give me the average energy that will be delivered for 1 clock time period for 1 clock cycle. If suppose this  $f_{sw} t$  will be nothing but 10 value. I will have to deliver  $10 CV_{dd}^2$  that is the energy delivered by the  $V_{dd}$ , it has to deliver 10 times.

But for one particular clock cycle what should be the energy delivered by the clock and that is dependent on  $\alpha CV_{dd}^2$ , that is given by  $\alpha CV_{dd}^2$  where  $\alpha$  is the activity factor that defines or characterizes the output node of the gate. This is what I have stated in the previous slide activity factor and the activity factor is generally either it is measured and then stated or you can also consider it to be a probability factor.

The probability that the current node, that is the transition from 0 to 1, because the  $V_{dd}$  supplies the energy only when the output changes from 0 to 1. It is basically the activity factor we can do two methods one method is to characterize the output node and see the history of it how many times it is going from 0 to 1 for a particular design and then while we are developing designing the new chip. If you are using the same subsystem or a system we can use that same characteristics.

The other way is to use the probability factor because the gates are also nothing but the logic designs. Based on the input combinations we can have a probabilistic method to identify how many number of times or what is the probability that the output node of the gate does the transition from 0 to 1. We are only interested in 0 to 1, because that is when the  $V_{dd}$  is going to deliver the energy or the power.

(Refer Slide Time: 21:36)

# Example: In 65 nm technology node,  $\lambda = 25 \text{ nm}$ ,  $f_{\text{clock}} = 1 \text{ GHz}$   
 A chip is designed which consists of  
 Number of Logic gates:  $50 \times 10^6$ ,  $W = 12 \lambda$ , activity factor = 0.1  
 Number of gates used in Memory:  $950 \times 10^6$ ,  $W = 4 \lambda$ , activity factor = 0.02  
 $C_a = 1 \text{ fF}/\mu\text{m}$ ,  $C_{diff} = 0.8 \text{ fF}/\mu\text{m}$   
 Estimate Power dissipated by  $V_{dd}$  for the Chip.

$P_{\text{avg}} = \alpha f_{\text{clock}} C V_{dd}^2 = \alpha C (f_{\text{clock}} V_{dd}^2)$   
 $P_{\text{avg}} = [\alpha_L C_L + \alpha_M C_M] f_{\text{clock}} V_{dd}^2$   
 $C_L = 50 \times 10^6 \times 12 \times 0.025 \mu\text{m} \times 1.8 \text{ fF}/\mu\text{m} = 27 \text{ nF}$   
 $C_M = 950 \times 10^6 \times 4 \times 0.025 \mu\text{m} \times 1.8 \text{ fF}/\mu\text{m} = 171 \text{ nF}$   
 $P_{\text{switching}} = (0.1 \times 27 \text{ nF} + 0.02 \times 171 \text{ nF}) \times 10^9 \times 1^2$   
 $P_{\text{switching}} = 6.12 \text{ W}$

NPTEL logo and a small photo of a person in the bottom right corner.

Moving ahead, I am taken an example here. In a 65nm technology node we know that the  $\lambda=25\text{nm}$ ,  $f_{\text{clock}}=1\text{Ghz}$  and generally this will be the case for a 65nm technology node. A

chip is designed which consist of many of the logic gates and then the memory also has many of gates.

The gates use this 50 millions here memory uses 950 millions. Total put together is I think 1000 millions. That width of the transistors which were used to design the logic gates was actually  $12\lambda$ , activity factor for this particular logic gates on an average is for  $50 \times 10^6$  logic gates is nothing but 0.1. The output of all these logic gates  $50 \times 10^6$  the average activity factor turns out to be 0.1, one could have a slightly better, one could have a slightly lesser value.

But overall, on an average it is considered to be 0.1 for the memory, the gates that has been used the width is  $4\lambda$  notice that w the width for the logic gate is  $12\lambda$  here the width for the memory is  $4\lambda$  here. Logic gates generally we need the computational to be faster, that is why we have used a higher width. Activity factor for the memory design is considered to be 0.02, it is again an average value. The memory whatever the logic gates 950 billion gates are used and an activity factor for on an average across 950 millions turns out to be 0.02.

The gate capacitance and the diffusion capacitance are given for these gates and it is 1fF per micron and  $0.8\text{fF}/\mu\text{m}$ . What we are supposed to find is estimate the power dissipated by the  $V_{\text{dd}}$  for the chip. This is basically what we are asking is an estimate the average power across the clock cycle. The average power by,

$$P_{\text{avg}} = \alpha f_{\text{clock}} C V_{\text{dd}}^2$$

The energy is nothing but  $\alpha C V_{\text{dd}}^2$  for one particular clock cycle. But if I want to find out the power it will be nothing but  $= \alpha f_{\text{clock}} C V_{\text{dd}}^2$ . This  $f_{\text{clock}} = 1\text{Ghz}$ ,  $V_{\text{dd}} = 1\text{ volt}$ .

What we will have is  $\alpha C$  and then this will be considered as a known parameter. This is basically known this is something we need to evaluate. The  $\alpha C$  we can have the  $\alpha$  parameters for the logic gates will be different and for the memory it is different. I am writing this as  $\alpha_L C_L$  for the capacitance of the logic gates, an activity factor for the logic gates and alpha for the memory design and C for the memory design.

$\alpha_L C_L$ , if I want to find it out the  $C_L$  is nothing but the number of gates here  $950 \times 10^6$  multiplied by each of these gates will have the gate capacitance and then the diffusion

capacitance. We are accommodating both the gate capacitance and then the diffusion capacitance for a width, this is per unit width of the transistor. Overall this capacitance is nothing but 1.8fF per unit width of the transistors.

$$P_{\text{avg}} = \frac{[\alpha_L C_L + \alpha_m C_m] f_{\text{clock}} V_{\text{dd}}^2}{V_{\text{dd}}}$$

$$C_L = 50 \times 10^6 \times 12 \times 0.025 \mu\text{m} \times 1.8 \text{fF}/\mu\text{m} = 27 \text{nF}$$

$$C_m = \frac{950 \times 10^6 \times 4 \times 0.025 \mu\text{m} \times 1.8 \text{fF}}{\mu\text{m}} = 171 \text{nF}$$

$$P_{\text{switching}} = (0.1 \times 27 \text{nF} + 0.02 \times 171 \text{nF}) \times 10^9 \times 1^2 = 6.21 \text{W}$$

Remember that this chip design it is still a very small chip design because it uses only the number of logic gates and then the memory and it is only 1, 1 particular die there could be multiple such dice or stack dice which are connected by the vias. If I have that kind of a 3D stacked one, the average power or the switching power or the dynamic power that has been delivered by the  $V_{\text{dd}}$  turns out to be 1 point 6.12 watts and slightly more than that.

If I have a similar one chip you consume 6.12 watts and if I have another chip, another chip and another chip and all of them are stacked together. Each one of this chip or each one of this die consume 6.12 watts then the overall design takes in more than 6 multiplied by 4 it turns out to be more than 24 watts which is very very high. In that sense what we need to do is what we have understood now is how to estimate the power, but after estimating the power we need to find some kind of a low power techniques. That the overall power we get some power benefits and then the overall power that has been delivered by the  $V_{\text{dd}}$  tends to reduce.