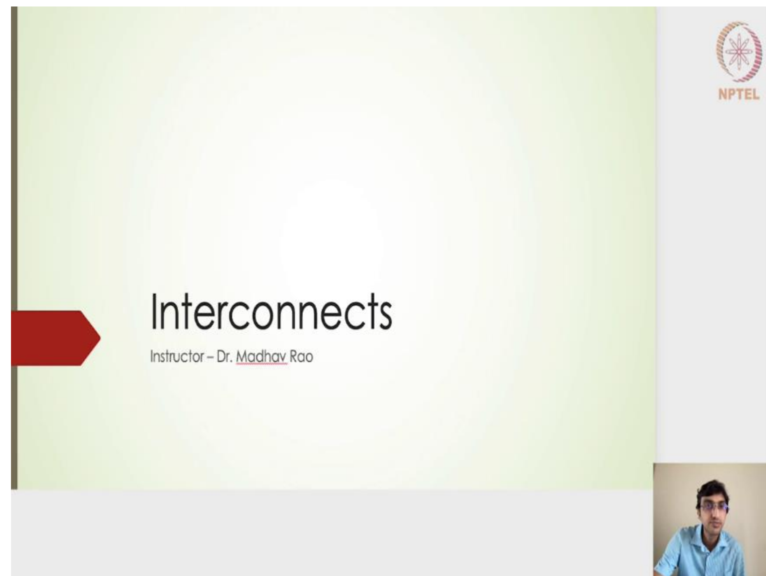
Design and Analysis of VLSI Subsystems
Dr. Madhav Rao
Department of Electronics and Communication Engineering
International Institute of Information Technology, Bangalore

Lecture - 53
Introduction to Interconnects

(Refer Slide Time: 00:16)



Hello students, welcome to this lecture on Interconnects. This is an introduction to the interconnects or also called as wiring in the CMOS VLSI design. Generally, what happens is whenever we are having lot of transistors the transistors are anyways connected by the wires to the other transistors. These are one type of interconnects the other type of interconnects are the wirings which connect from one subsystem to another subsystem.

In that sense if one is actually designing a subsystem and then in one part of the chip and then there is another subsystem which is there on the other part of the chip and there will be both the systems has to be connected. Then a wire an electrical wiring has to be done from one end to the other end and that too on the chip, that forms another type of interconnects.

There are actually several types of interconnects the global level interconnects or the local interconnects and so on. That is where the interconnects play plays a very crucial role in the overall system design in the chip. What we are going to learn in this particular lecture

series about interconnects is how the interconnects or the wiring is going to affect the overall performance of the design.

If I have a very simple form if I have an inverter connected to an another inverter and if I see a kind of a long wire what is bound to happen, whether the inverter connected to an inverter without the interconnects how does it perform and then with the interconnects how does it perform.

(Refer Slide Time: 02:08)

The basic characteristics of the interconnects is in the form of a resistance or a capacitance. Generally whenever a wire element or a interconnect component is available in our design we try to characterize that in terms of the resistance and then the capacitance.

Once we have the resistance and then the capacitance we can actually integrate with that of our transistor resistance and capacitance model and then we can use any kind of the Elmore delay method to estimate the overall delay or to estimate any other parameters.

Generally you will see that the interconnects are kind of characterized in terms of the resistance and then the capacitance. Now, how do we evaluate a resistance and the capacitance given an interconnect? It is very the resistance part of it is quite simple it is kind of very elementary. If you look into this particular wire A and B and I have shown here as a block of wire with the width of w and then the length of l for this particular wire and then there is a thickness t.

As per our Ohm's law or the definition the resistance is nothing but of course the voltage by the current, but in terms of the material or the dimension level definition of the resistance it turns out to be,

$$R_{AB} = \frac{\rho l}{A}$$

Where ρ is the resistivity of the material that has been used and l is the length dimensions and then area = w x t.

This can also be rewritten in the form of,

$$R_{AB} = \frac{\rho}{t}\frac{l}{w}$$

$$R = R_{\square}\frac{l}{w}$$

We can also express it in the form of the sheet resistance.

$$R = R_{\square}\frac{l}{w}$$

Now, this sheet resistance has a very important property that the reason is this the thickness. The thickness here is actually a technology or a process parameter we cannot change with respect to the design. It is a process dependent and rho is a material dependent again we can call it to be a process kind of a dependent. For example, 65nm technology or 180nm technology uses the copper as the wiring material or the interconnect material.

This remains fixed for a process parameter and then t remains fixed the thickness is usually comes from the deposition process and for all the interconnects at different levels we will define the thickness. In that sense the thickness is kind of defined for every layer or for every level in a particular process technology node. If I пck one level of the interconnects then I its thickness is defined or it is know we can consider it to be constant.

In that sense we can redefine this rho by t as its sheet resistance and if we know this sheet resistance for a particular process node with different levels then we can easily calculate the resistance of the interconnect at different levels based on its design parameters l and

w. The units of the sheet resistance is pretty simple it is nothing but this is a resistance the unit is Ω. The sheet resistance will have actually Ω and Ω/square alright.

The copper has a resistivity of,

$$\rho_{cu} = 2.2\ \mu\Omega cm$$

The aluminum has a resistivity of,

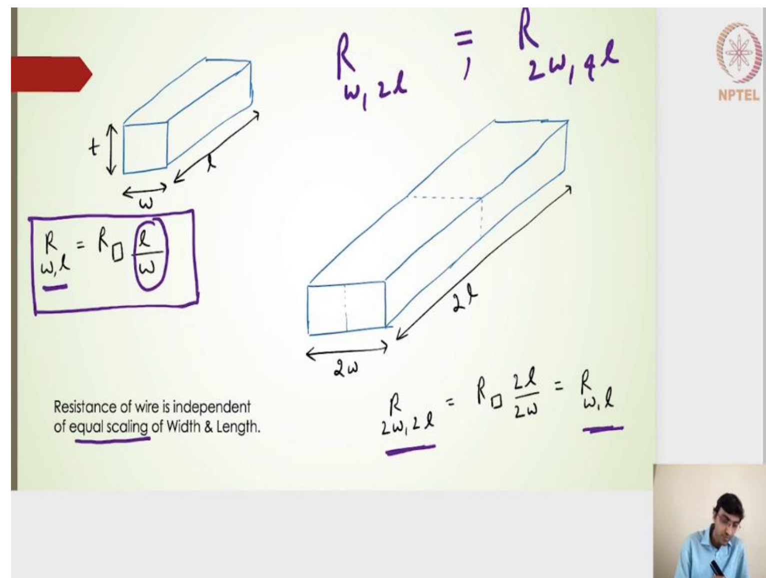
$$\rho_{Al} = 2.8\mu\Omega cm$$

Aluminum is generally kind of outdated and it is mostly copper which is being used in the recent semiconductor industry or in the regular process parameters. You will see mostly the copper as the wiring material in most of the process technology nodes and aluminum will not be used.

But, I have given both the aluminum and copper just to look at this values of 2.2 and 2.8 which implies that the copper has a better conductivity properties in the sense that the resistivity is low for the copper when compared with that of the aluminum and in that sense the conductivity of the copper is better and what it really helps is if the conductivity is better the current will be better and thereby it improves the performance.

The other advantage of having the copper is, there is standardized copper reverse electroplating method and that has become more standardized that is been very well characterized or very well calibrated and in that sense the copper for the deposition the copper deposition process has become really a well known or an accepted method. That is one of the reason why the copper has been are showing good results and it is widely accepted.

(Refer Slide Time: 07:48)



Moving ahead this is what we have for a dimensions of w and l of the copper wire we have,

$$R_{w,l} = R_{\square}\frac{l}{w}$$

Now, tomorrow if I change this dimensions to 2l and then 2w. This is what I have shown here

$$R_{2w,2l} = R_{\square}\frac{2l}{2w} = R_{w,l}$$

Assuming it is the same level in the process in the technology node process then the thickness will also be the same only thing that will change is w and l. If it increases twice that of the first block of wire still the resistance remains the same because the l and w which is taken as a ratio here it will be scaled by 2 times, but 2 and 2 will get cancelled and then finally, we will have a resistance equal to that of w,l.

Resistance of 2w,2l will be similar to that of the resistance of w,l. The resistance of the wire is actually independent of equal scaling of the width and the length. What it really implies is if I have l and w and if I equally scale it both the length and the width to 2 times or to the n times I will once again get the same resistance alright.

Now, if suppose I have a resistance of say Rw,2l and tomorrow I will increase it to say 2w,4l then also this both this resistance is equal. Equal scaling on the length and width will give me the same resistance right this is clear.

(Refer Slide Time: 09:37)

Moving ahead, just an example to find out what should be the resistance of the copper wire, generally we will see in the VLSI design. The thickness parameters for a 65nm technology known for one of the medium levels will be 0.22μm that will be the thickness, the width for this particular design is 0.125μm and suppose if I have a long wire of 1 millimeter. What should be the resistance and also wanted to evaluate what should be the sheet resistance?

The sheet resistance is pretty simple based on the resistivity of the copper and then divided by the thickness which is 0.12μm or rather the thickness is 0.22μm.

$$R_{cu} = \frac{2.2\mu\Omega cm \text{ x } 10^{-1} cm}{0.22 \text{ x } 10^{-4} \text{ x } 0.125 \text{ x } 10^{-4}} = 800\Omega$$

$$R_{\square,cu} = \frac{2.2\mu\Omega cm}{0.22\mu m} = 0.1\Omega/\square$$

What it means is with this particular the dimensions which we will normally see in the digital VLSI design on a chip the resistance is around 800Ω and this will have an impact on the overall performance of the design.

(Refer Slide Time: 12:00)

As I said earlier the interconnects will always be characterized with the resistance and then the capacitance. How do we evaluate the capacitance? one of the very simple method to estimate the capacitance of the wire is using the parallel plate capacitor model.

What it means is if I have a interconnect or a wire component here with the dimensions of wl and then the distance between the another metal wire it could be running horizontally or it could be running laterally. In that sense if the distance between or the spacing between them is h here then I can easily calculate what is the capacitance here. Because we can assume this to be one metal plate the another one I have written as a ground plane it could be any other metal layers beneath it or above this particular interconnect or the wiring component, anything above this and then below this will form kind of a parallel plate capacitor.

Because we have two metal plates one here one of this one and then another one beneath this or it could be above this also. That forms the second metal plate and between that we have the dielectric it is always isolated by the dielectric silicon dioxide. We get this capacitor there between the two plates and then sandwiching the dielectric material.

In that sense the parallel plate capacitance value is nothing but,

$$C_p = \frac{\varepsilon_{ox} wl}{h}$$

$$\text{Where, } \varepsilon_{ox} = \varepsilon_o K_{ox}$$

$$K_{ox} = 3.9$$

$$\varepsilon_o = 8.854 \text{ x } 10^{-14} \text{F/cm}$$

The w and l parameter if it is known and then the spacing h is known then we should be able to find out what is the capacitance that this particular interconnect is characterized.

Now, remember that the,

$$C_p \alpha l$$

$$R_{wire} \alpha l$$

In that sense both of them is directly proportional to the length of the wire. If I keep increasing this length of the wire both the capacitance and then the resistance of the wire is likely to increase.

It will have a drastic effect on the overall system performance. The other parameter to be interesting to note is the w parameter.

$$R_{wire} \ \alpha \frac{1}{w}$$

$$C_{wire} \ \alpha \ w$$

An increase in the w here is likely to decrease the resistance of the wire, an increase in the w here is likely to increase the capacitance of the wire. Turns out to be very very similar relationship when we calculated or when we tried to model the transistors with the width. When we try to scale the width of the transistors its resistance used to decrease, but its parasitic capacitance used to increase.

In the similar sense we will have this particular parameter and somewhere this will be form a major component in optimizing the design width of the wire, similar to how we had seen the transistor width and how an optimized width is required for the transistor so that we get the minimum delay.

Similarly, the same ideology will be used to identify the optimized width of the wire. That we will get the minimum performance or the minimum rather minimum delay for the overall circuits which consists of the interconnects. So, hope this is clear.

(Refer Slide Time: 16:23)

.

Let us look into the modeling the delay for an interconnect. Let us say that the interconnect is 1 mm long. The 1 mm long interconnect or one can also argue that it could be a very long wire. So, as such if I have a very long wire we can always have it as a distributed component and then use that particular distributed component to model our interconnects resistance and then the capacitance and then use that or integrate that into our circuit for estimating the overall performance delay power and other aspects.

What we are trying to see here is interconnect is a distributed circuit in the sense if I have a long interconnect A to B we say that if I can actually distribute it among different segments each of these segments will have a resistance will have a capacitance, each of the segments will have a resistance capacitance and each of the segments will have a resistance and capacitance.

That is what the distributed circuit says where R and C are modeled as the lumped elements in the sense the passive elements resistance and then the capacitance. If you do this particular distributed model of an interconnect, it actually becomes a very simple it simplifies our overall estimation of the delay and power and other statistics.

But, while we are actually distributing the circuit. Each of these segments could be modeled as an L model, it could be modeled as a π model and it could be modeled as the T model and the naming of this particular model is actually the way the R and C are positioned.

The L model will have the R and C like this the π model will have C by 2 component and then in between that there will be a resistance R, the T model we have R/2 and R/2 and then there will be a capacitance in between that. Notice that in all the 3 models we are considering single segment or a single distributed part of the interconnect where the single segment is having an overall resistance of R and a capacitance of C.

Similarly in the π model the overall segment should have a resistance of R capacitance of C. That is why being a π model it is actually showing C/2 here on one side and C/2 on the other side. Then similarly T model one segment each segment has a total resistance of R and then the capacitance of C. In the T model what it does is R/2 it separates it out on one side and the other side and in between there is a C.

If I want to find out the overall delay here for this particular L model circuit it turns out to be,

$$\mathrm{delay} = \mathrm{RC}$$

For this one if this is the input signal and then this is the output I will get,

$$\mathrm{delay} = \frac{\mathrm{RC}}{2}$$

For the T model if this is the input and then this will be the output because,

$$\mathrm{delay} = \frac{\mathrm{RC}}{2}$$

Even if this one and then if I take this one both the outputs will be similar there will not be any current that will be flowing here and that is why whatever is the voltage that has been accumulated here it will be the same as here.

In that sense the delay will be nothing but R/2 and multiplied by C, it will be R C/2. Notice that for the L model I will get an delay of RC for π by 2 π model it is the R C/2 and then for the T model it is R C/2.

(Refer Slide Time: 20:26)

Let us go to the three segments. I have an interconnect I have a wire here basically its resistance is R and its capacitance is C and now we have actually distributed into 3 components, the 3 parts. This is one part 1 segment, this is 2nd segment, this is the 3rd segment.

What we are saying is it is been distributed. The overall resistance is actually distributed to R/3, C/3 this one is also distributed to R/3, C/3 and then this one is also distributed to R/3, C/3. Overall I will have a resistance of R and a capacitance of C/3.

If I consider this particular model or this particular segment in the L model. If I consider this particular segment the 1st segment 1st of the 3 segments and then actually use the L model for the 1st segment I will get R and C. This is basically R/3 and C/3, 2nd segment is this one R/3 and C/3, then the 3rd segment is R/3, C/3.

The overall delay for the 3 segments will be nothing but starting from here and then here. It will be using the Elmore delay method I should be able to find out what is the overall delay.

$$\text{delay} = \frac{R}{3}\frac{C}{3} + \frac{2R}{3}\frac{C}{3} + \frac{3R}{3}\frac{C}{3}$$

$$\text{delay} = \frac{RC}{9} + \frac{2RC}{9} + \frac{3RC}{9}$$

$$\text{delay} = \frac{2}{3}RC$$

If I use the single segment in the form of a π model then I will get C/3 and R/3. I will have C/3 two parts I will do two parts of C/3. I will get C/6 on this side C/6 on this side and in between I will R/3.

This will represent the 1st segment, the 2nd segment will be the similar C/6, C/6 on either side of R/3 and then the 3rd segment will be C/6, C/6 and then R/3 in between that. Overall I will get a delay of again RC/2 here. If I do R/3 into C/6 and C/6 in parallel will give me C/9, C/3 here.

$$\text{delay} = \frac{RC}{9} + \frac{2R}{3}\frac{C}{3} + \frac{3RC}{18}$$

$$\text{delay} = \frac{RC}{9} + \frac{2RC}{9} + \frac{3RC}{18}$$

$$\text{delay} = \frac{RC}{2}$$

For the T model individual segments is having a resistance of R/3 and C/3. Here in the T model the resistance is actually divided into two parts.

One part it is kept on one side of this C/3 the other side it is the remaining resistance. It becomes R/6 and R/6 and in between there is a C/3 that this will be one segment, the similarly 2nd segment will be this one, 3rd segment will be this one and the overall delay will be nothing but,

$$\text{delay} = \frac{RC}{18} + \frac{RC}{6} + \frac{5RC}{18}$$

$$\text{delay} = \frac{RC}{3} + \frac{RC}{6}$$

$$\text{delay} = \frac{\text{RC}}{2}$$

Notice that in the previous slide when we had only 1 segment this particular interconnect was actually seen as 1 segment we got a delay of R C in the L model, we got a delay of RC/2, we got a delay of R C/2 in the T and п model whereas, if I make it into 3 segments now the delay is for the L model is actually 2/3 RC which is actually reduced from RC to 2/3 RC.

The п model it remain the same RC/2 and the T model also remain the same RC/2 and if I actually make 10 segments for the L model, п model and T model and if I do the 10 segments.

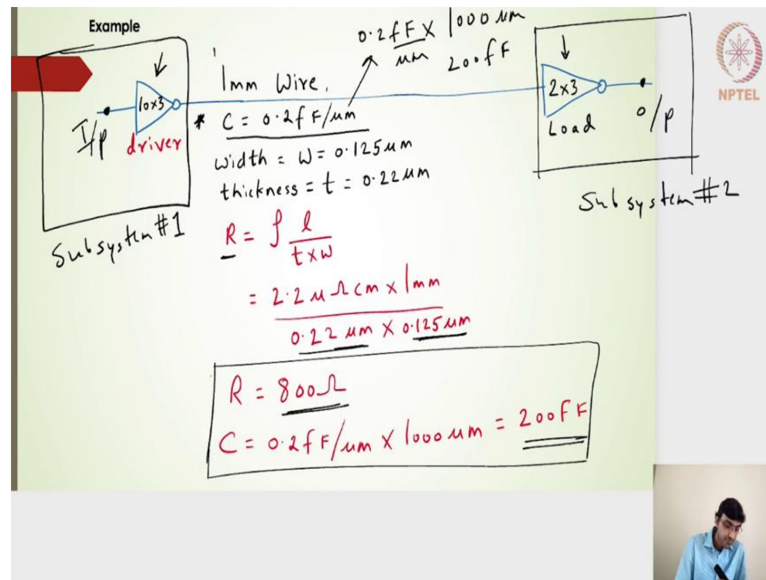
(Refer Slide Time: 24:22)

I will actually get RC/2 here almost close to R C/2 for the L model, it is actually decreasing. If I had 1 segment of the wire with a characterized with the resistance of R and C and if I use L model I will get the delay as RC, if I make 3 models I will actually get a decreased delay of RC/3. Finally, if I make it 10 segments I will get close to RC/2. For the п model and T model it was giving me RC/ 2 in the 1 segment it was giving RC/2 in the 3 segments.

In the 10 segments also it is giving me the value of RC/2 same here for the п as well as the T model alright. What it implies is, if I have an interconnects either use the п model which is more standard or the T model, both of them gives us the correct estimate of the delay.

The delay of the interconnect, characterized with the resistance of R and capacitance of C actually gives us a delay close to RC/2. This is a more appropriate or a well accepted model alright. Here in this particular 10 segments the calculation is done in a way that each of the segments has R/2 the distributed resistance and then the capacitance alright.

(Refer Slide Time: 26:01)



Moving ahead, this is one such example I wanted to give. I have a driver inverter and the driver inverter is connected by a long wire to the load inverter. What it means is I have an inverter here or it could be one part of a subsystem 1 design.

Subsystem module number 1 and from here onwards we can say that this could be the another subsystem design alright. Between this subsystem 1 and subsystem 2 there is a electrical wiring or there is a basically an interconnect that is connecting the subsystem 1 and subsystem 2. If I want to see the delay from the input here to the output here which is connected by a long interconnect, then what should be that particular delay or the performance.

In this particular lecture we will just try to see what is the overall capacitance or the resistance associated with this particular interconnect as well as the driver capacitance and then as well as the load capacitance. The driver resistance switching resistance and then the parasitic capacitance and then the load input capacitance as well as the switching resistance.

Before that let us try to evaluate what should be the overall resistance and then the capacitance of this long interconnect. This particular long interconnect wire =1 mm, width w = 0.125µm and thickness t = 0.22µm, C = 0.2fF/µm.

If it is 1 mm wire the overall capacitance of this particular wire will be nothing but 0.2fF x 1 mm wire, I will have to multiplied by 1000 µm. This will be 0.2fF x 1000µm. I will actually get somewhere around 200fF.

This is what it is 200fF is the overall capacitance of this long interconnect. The interconnect is actually characterized with the dimensions given 1mm width and thickness and it is also characterized with the capacitance per unit length of the micron as 0.2fF/µm. The resistance of this particular wire is nothing but 800 that is what we had calculated in the previous slide.

$$R = \rho \frac{l}{t x w}$$

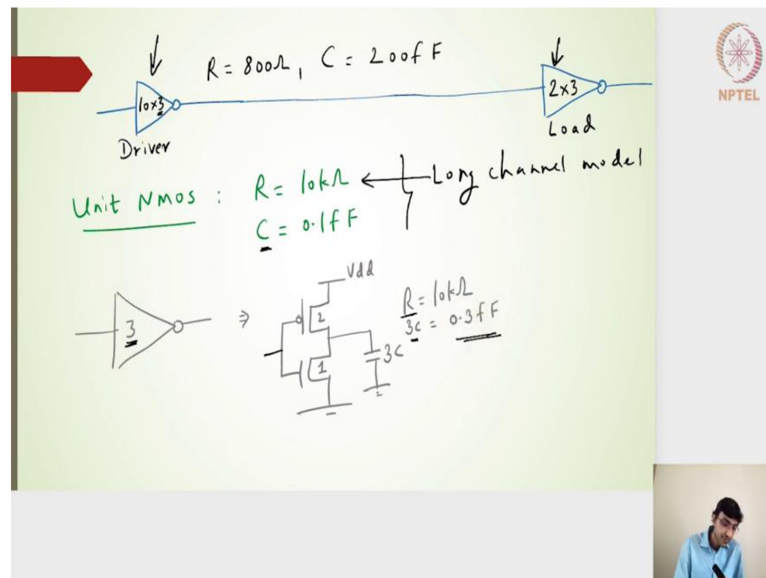
$$= \frac{2.2\mu\Omega cm \text{ x } 1mm}{0.22um \text{ x } 0.125\mu m}$$

$$R = 800\Omega$$

$$C = 0.2fF/\mu m \text{ x } 1000\mu m \ = 200fF$$

I now calculated the resistance of this wire and then the capacitance of this wire which is 200fF and the resistance is 800Ω. Let us try to understand what should be the switching resistance and then the parasitic capacitance of this driver and what should be the input capacitance of this particular loaded inverter alright.

(Refer Slide Time: 29:34)



Let us come back to this unit NMOS. The unit NMOS transistor will give me a switching resistance of 10kΩ. Assuming this is a long channel model, the switching resistance we got was around 10KΩ, the capacitance was on an approximate value we used to take 0.1fF. If I have a unit inverter of 2:1 ratio, the input capacitance at C is a 3C capacitance, the parasitic is 3C here I used to get a 3C value as 0.3fF because unit NMOS transistors capacitance is 0.1fF.

The overall capacitance should be 3 x 0.1fF = is 0.3fF. The switching resistance for a size of 3, 2:1 which is nothing but the size of the 3 will give me 10KΩ. In that sense if the driver is 30 that is 10 x 3. The 10 times the unit, 10 times the gate size of this particular driver inverter is 10 times more than that of the 2:1 inverter.

In that sense my resistance the switching resistance is likely to decrease by 10 times and my capacitance instead of 0.3 it was going to have multiplied by 10 times. Similarly on the load side if it is 2 times more than that of the gate size is 2 times more than that of the unit inverters 2:1 inverters size then the resistance the switching resistance will get decreased by 2. The capacitance is going to increase by twice, the capacitance will be instead of 0.3 it will be 0.6 for the load 1 and this will be 0.3 x 10 for the driver 1.

(Refer Slide Time: 31:34)

For the driver 10 x 3, the switching resistance of the driver turns out to be,

$$R_{driver} = \frac{10K\Omega}{10} = 1K\Omega$$

The switching resistance of the load will be nothing but,

$$R_{load} = \frac{10K\Omega}{2} = 5K\Omega$$

$$C_{driver} = 30 \text{ x } 0.1fF = 3fF$$

$$C_{load} = 6 \text{ x } 0.1fF = 0.6fF$$