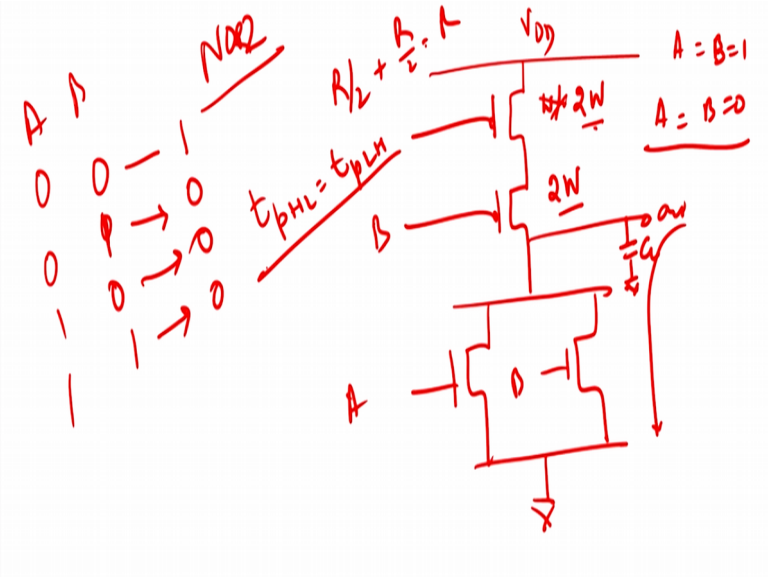
CMOS Digital VLSI Design
Prof. Sudeb Dasgupta
Department of Electronics and Communication Engineering
Indian Institute of Technology – Roorkee

Module No # 03
Lecture No # 14
Combinational Logic Design –III

Welcome to the NPTEL online certification course on digital VLSI design and this combinational logical 3 will start today. And let us see what you discussed till now we will see in that propagation delay and noise margin are input dependent. So we see in that for tow input NAND gate you want that your width of your NOMS which is the pull down should be double that of PMOS in order to ensure that tpHL and tpLH are equal. Let us see as I left the previous section how does it work out if you have a two input NOR gate.

(Refer Slide Time: 01:03)

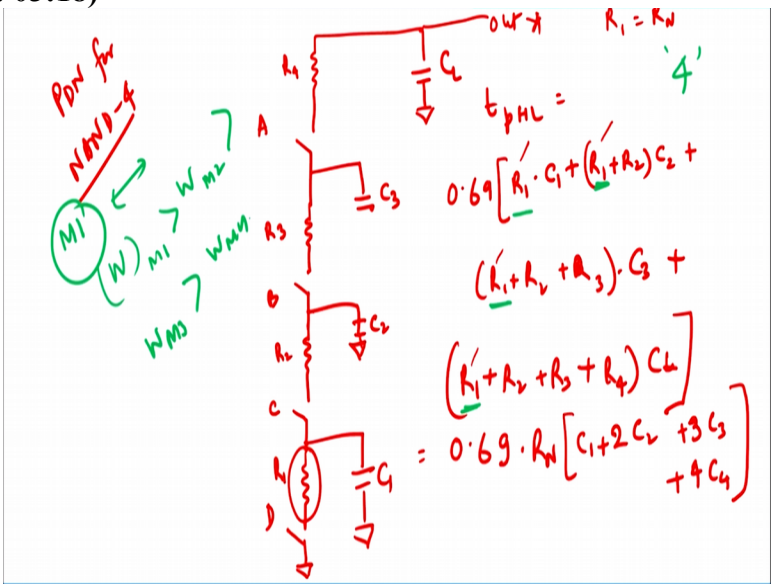


Then what you what are our comments as far as pull down and pull up network relative sizing is concerned in terms of aspect ratio so this is VDD and we had A here and we have B here we have A here and we have also B here right now if you look at this is your output here so this is basically your NOR2 logic and as I discussed with you in the previous case you have this output if this is A and B output will be always equals to 1 here and 0, 1 will give you 1 and therefore you will always be the 0.

Now you see in this case just the reverse trend will happen that for pull down you do not have to make both A and B equals to 1. If you make either of B = 1 your job will be doe and the output will go round this capacitance CL will actually low to ground voltage whereas for pull up now pull up now both A and B have to be 0 and therefore now W / L ratio of this one for making up and down same W/ L ratios for PMOS should be actually double that the first part.

So this should be 2W this should be 2W right so for a non-logic is just as the reverse you want the PMOS's to care because there will be series to each other. So make the resistance R by 2. R / 2 + R / 2 will give you + R / 2 will give you equals to R if you assume that the electron mobility (()) (02:43) are equal this is the basic assumption then we assume that the pull down capability of a single PMOS is exactly good to that of NMOS then in that case you just have to widen the value of the first transistor and that the pull up transistor once you do that you end up having the resistances half and therefore the low to high transitions are equals to high to low transitions therefore I get tpHL will be equals to tpLH and that is what is quiet interesting which we get.

(Refer Side Time: 03:18)



So if I get four input NAND gate for example and can draw its corresponding diagram in this manner there I take four input NAND gate I am trying switch capacitor or a switch layer design so this is R4 right and then you have got S1 here let us suppose this is A then we get R3 here and we get here R3 let us suppose this is B and then I get R2 here right and then I get C and then I get R1 here and then I get ground here and then C this is R1 and individual CL this will be your C3 this will your C2 right and this will be of course will be your C1 right.

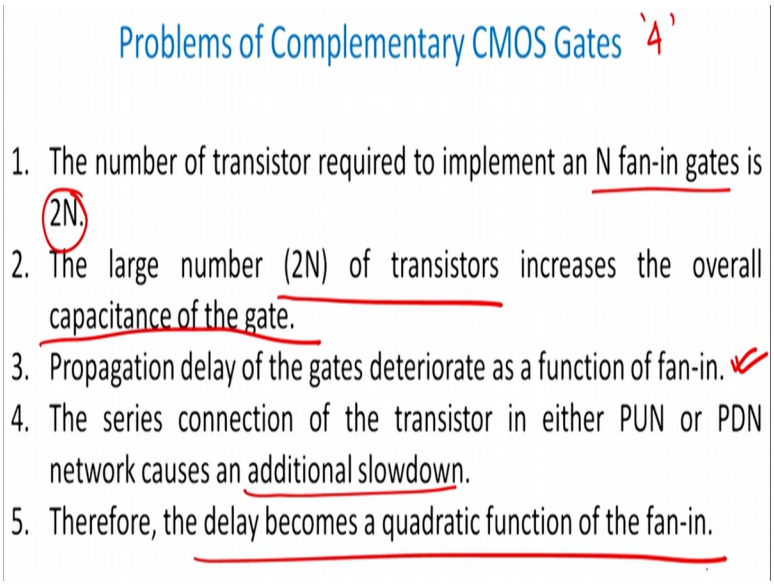
And this is your load capacitance in the output side so this is the pull down network PDN for NAND 4 and you can see this is the overall scheme. So if you want to find out the value of tpHL how to low it comes something like this right. So you see all the four transistors have to switch on in order to generate the logic or generate the high to low fall. So for doing that I get 0.69 right multiplied by R1 times C1 + R1 + R2 times C2 right + R1 + R2 + R3 times C3 right + R+ R2 + R3 + R4 times CL right CL this is your overall delay.

Assuming that so this will come out to be tpHL will out to be 0.69 times if you take all the resistance to be equal and assume it to be 1. So R1 = RN so I get RN in the outside the bracket and I get C1 + 2C2 + 3C3 + 4 CL right I get this into consideration quiet an interesting part if you look at this point therefore that R1 is coming four times to you 1, 2, 3 and 4 right. So if you want to optimize this speed your NAND four logic using complementary MOS technology then please keep in mind that your try to optimize the lowest one the one closest to the ground.

If you try to optimize that you will get the fastest delay actually optimize at a much higher rate as compared to anything else which means that it is always preferable it is always preferable to have the optimization done at a transistor which is furthest away from the output node right. So the output node is this one in front of you and we need to optimize the MOS transistor which is furthest away is there in the R1 in this case right.

And this is what I just want to give your output or give a feedback hat generally those transistors you should be able to handle which are furthest away from the. So with this idea we have already understood this whole logic let me come to the next case what are the problem areas of complementary logic.

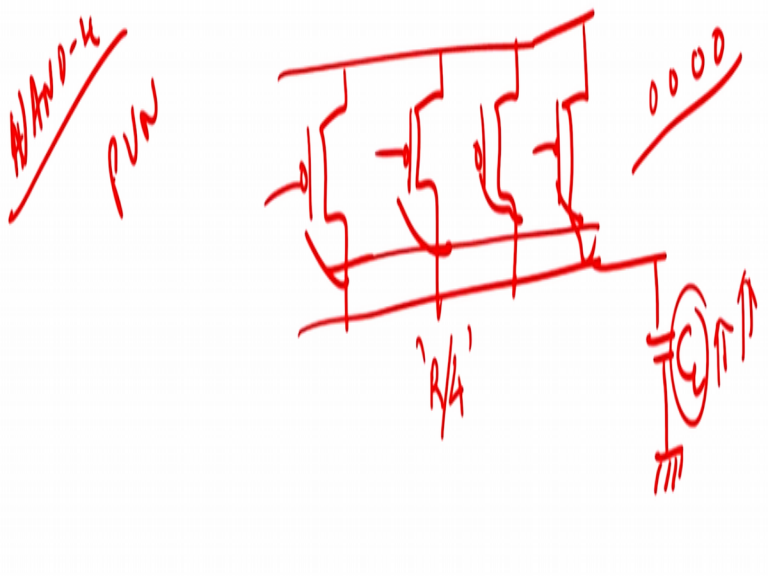
(Refer Slide Time: 06:52)



As you can see the problem area is basically two are three points the first point is that if you want to therefore generate and N input gate or N fan in gate you require two N transistors. So if you want that R2 logic and NAND 2 logic you actually require transistor to do it right and therefore your power dissipation levels are relatively higher also if silicon layout area using is also very larger.

So in both the cases this is the bottle neck available to you therefore if you have larger number of transistors your capacitance actually increases of the gate the capacitance actually larger why? Because if the number of transistor increases the number of overlap capacitances the number of Cox the number of CGD, CGS which we have discussed in the previous turn will also increase and therefore the overall cap loading capacitive loading will be larger right and therefore that will result in what result in much larger for example delay.

(Refer Slide Time: 08:05)



As I discussed with you therefore in a previous discussion if we have a 4 input NAND gate and the same thing if we have drawn for 4 input NAND only 4 NAND 4 we define this is NAND 4 but we now look at the pull up network. Pull up network if you remember is just nothing but 4 parallel PMOS transistors right but you see if you take NAND 5 now there will CL which is primarily coming because of what because of gate to drain because of which you have a CL capacitance.

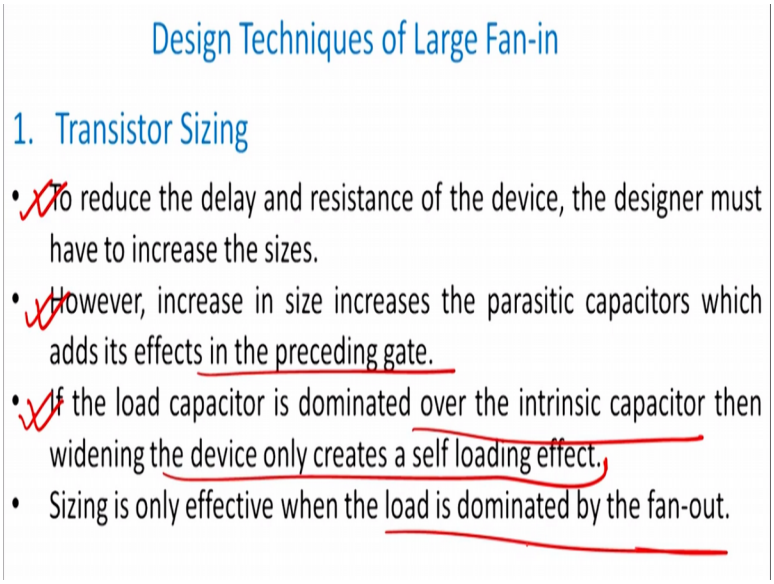
Now if you reach the fan in and make it work 5 input NOR then there will be further increase in the value of CL and therefore though in the condition when 0, 0 4 input you are using I will get a very fast charging of CL but it will be partly offset by a larger value of C itself are you getting my point. So that is a trick the trick is you have to therefore optimize between the fan in which means that the number of active device which was falling on the value of CL and this speed itself the value of CL itself.

So as you increase the number of PMOS in the pull up network right the overall resistance falls down say R / 4 right but in the same instance of time your also adding CL here right and therefore the CL value goes on increasing right then as a result the gain which you get by virtual lower resistance is slightly off set by a larger value of CL. So we need to find out the methodologies which you can have a minimum value of typically use no more than 4 as you have fan in.

For most of the purposes we use four fan in the maximum value in current digital VLSI system as discussed with you just now that the propagation delay of the gates deteriorate as the function of fan in so higher the fan in higher will be the deteriorate because you have a larger drop available to you. Further as I discussed with you just now that if you PU in case of NOR2 in case of NAND 2 logic P pull down network consisting of NMOS and series and in case of NOR logic pull network consisting of PMOS in series in either two cases higher the fan in more will be resistance of the path available between output and ground or VDD and therefore larger will be the delay available to you such a criteria.

So the delay is therefore additional slow down available to in this case clear people have seen to the details of quadratic function of an which means that you increase it by twice and our times the delay right and that is what the problem area which you see because in most practical cases it is R in to CL so CL increases drastically and therefore you delay in huge drastically let us look at what are so which is now see the saw that having a large fan in does not help my case in reduced tpHL or tpLH in fact my time period becomes larger and larger right we have also seen that this tpHL or LH is a direct function of the input voltage the noise margins are also direct function of input voltage.

(Refer Slide Time: 11:23)



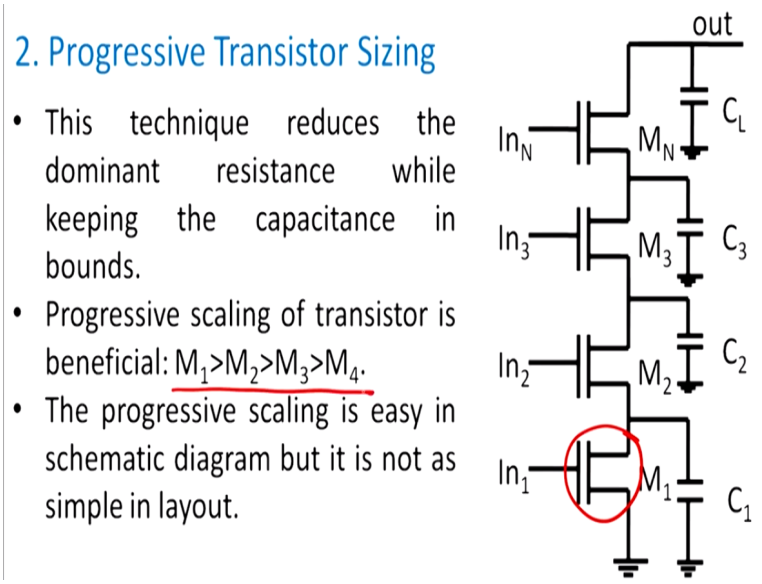
This is our input values of data right so where the data is varying very fast will get some value of tpHL and other values of tpLH. If you see that you are ending up having very large delay the first thing which you learn is basically try to increase the size of transistor right try to increase the

size of the transistor. As you start increasing the size of the transistor obviously the resistant's falls down and therefore your delay starts to fall down.

However has therefore the second point is always mentioned however if you increase your size you also increase the parasitic capacitance and which adds to the overall capacitances in the system. So simply by increase this size when lot make a life easier but at same instant of time intrinsic delay may go on increasing and therefore that will result in problem area. We have also learnt that if the if the load capacitance is dominated over the intrinsic capacitors what is the load capacitance CL which is consisting of subsequent stage input gate capacitances and the previous stage CGD and CGS CGD for NMOS and PMOS

That is the load capacitor CL much larger intrinsic capacitance of a devices then widening the devices only creates a self-loading effects we also seen this point earlier but you increase the value of W then it is only going add to the value of CL which you see in front of you right this is now self-loading phenomena or self-loading effect. As a result we will start to go up become larger and larger in each case as I discuss with you therefore the load is dominated by the fan-out higher the value of fan-out higher will be the value of CL well needless to say that why is it like that right now the methodologies which people have been using for delay optimization is that.

(Refer Slide Time: 13:30)



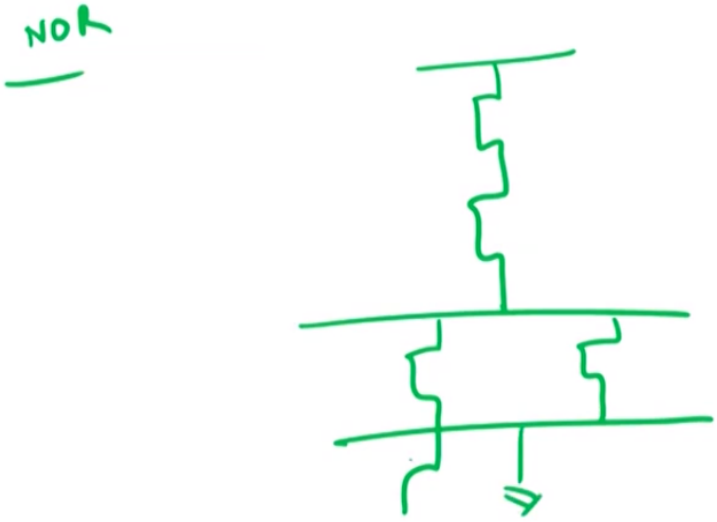
As I discuss with you just now that M1 will be coming from 4 times in the whole network right we saw we just now saw it that if you go for here if you look at R1 and this R1 is coming from

here as well as here this coming 4 times right. So methodologies chosen that if I able to optimize the value of M1 why because M1 is the transistor which is coming 4 times in all these respects.

If I am able to my W of M1 larger as compare to W of M2 larger than the W of M3 and larger than W of M4 then I will be able to optimize it much better manner. Why because since R coming 4 times so if you reduce R1 by even 10% there will be a contributory or colliding effort and then overall delay will be reduced drastically right. So the methodologies people adopted and has been used required long time is that the progressive scaling of transistor is beneficial right very important.

So those transistor which are further away from the output source in the pull down network if you size its larger you get this type of network right. Again since you understood for NAND 4 let me explain to NAND let us say NAND 2 may be yes NAND 2 let say see or NOR2 may be.

(Refer Slide Time: 14:59)



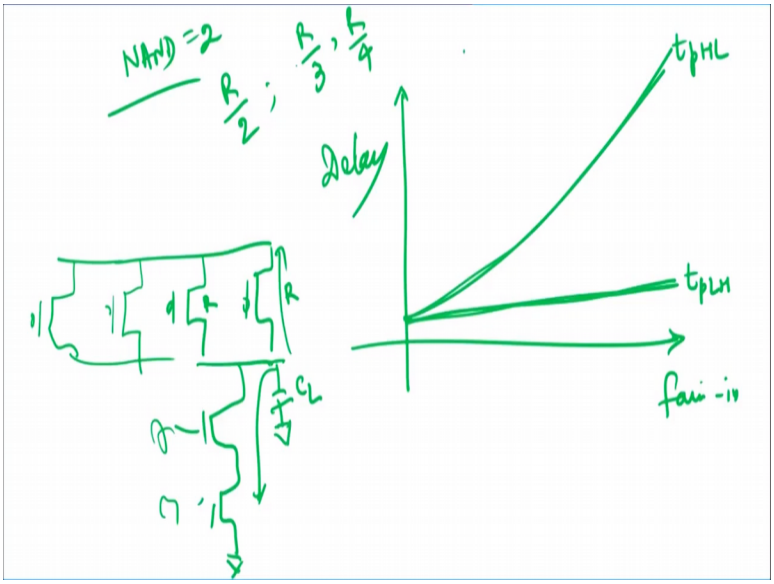
NOR 2 will be easier for you to understand in NOR 2 unlike in NAND you will have pull up which is basically as I told to you will be in serious and the pull down will be in parallel right. So I have got this VA. In this case however tpHL output side will be governed by either of the two transistors whereas tpLH low to high will be governed by this transistors low to high right now you see this transistors if it would have been a 4 input NOR then if it is basically NOR 2 for NOR 4 if you have got NOR 4 then I will have 1, 2, 3, 4 are available here and then you will

have this is your pull up network and then you will have similarly pull down network available here.

Now if this is what you get right then try to optimize this M1, M2 let us suppose M3 and M4 try to optimize the M1 in this the M1 is 4 time at least when you are calculating tpHL low to high right. So the size in criteria which we have already understood is that the propagation you should always do a progressive sizing which primarily means that your sizing should be going on increasing as you move going on decreasing as you move from log from the furthest part of the output towards the output.

This we already seen and we have explain to you as you can see here but R1 here will be the largest followed by R2 followed by R3 and followed by R4 right this technique is known as progressive scaling or progressive transistor sizing right as must transistor between use long duration of time in many of cases. Now let me give you an example and let see how it works just before doing that another basic example and see how it works out.

(Refer Slide Time: 17:07)



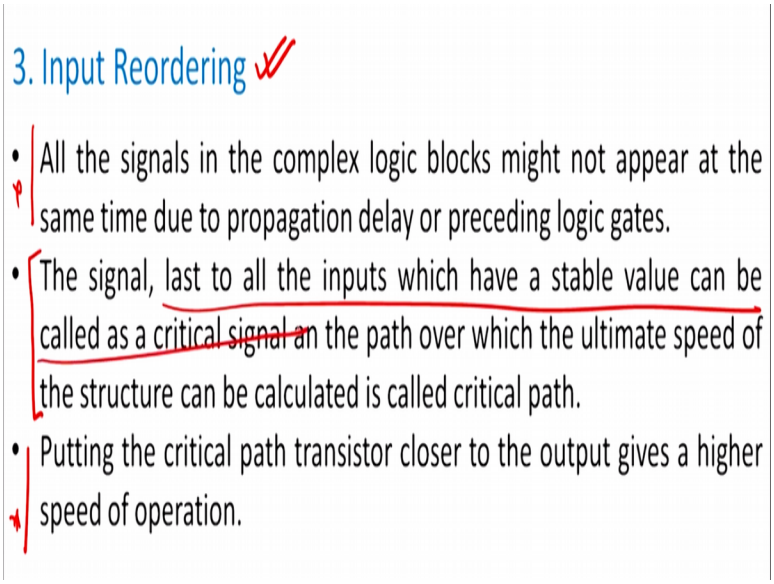
That let us suppose I have a NAND gate right I try to plot your fan-in and I try to find out the value of tpHL and tpLH then this is how it looks like. This is tpLH and this is tpHL this is for NAND 2 logic right and this is the variation of so this is your delay and with fan-in NAND if you remember was basically your pull up was basically your 2 PMOS is parallel and this is 2NMOS is in series right.

So tpLH low to high means this one is what I was taking about low to high this is grounded this is A and this is B this is grounded now if you see tpLH right tp low to high is almost independent of that fan-in you can understand why is it like that because as the fan-in increases the number of PMOS in the pull up network also goes on increasing yes it does act to act CL but it also lowers the effective R when you have a two input NAND right.

My so this is assuming that each one is R here when you have two inputs here two input NAND then you have got R / 2 right we I got a three input NAND available to me I will get R by 3 when I have 4 input NAND I get R / 4 and so on and hence so forth which means that higher those seen increasing slightly may be but reduction is R = R1 / R2 so there also balancing each other and therefore my tpLH low to high is almost remaining constant.

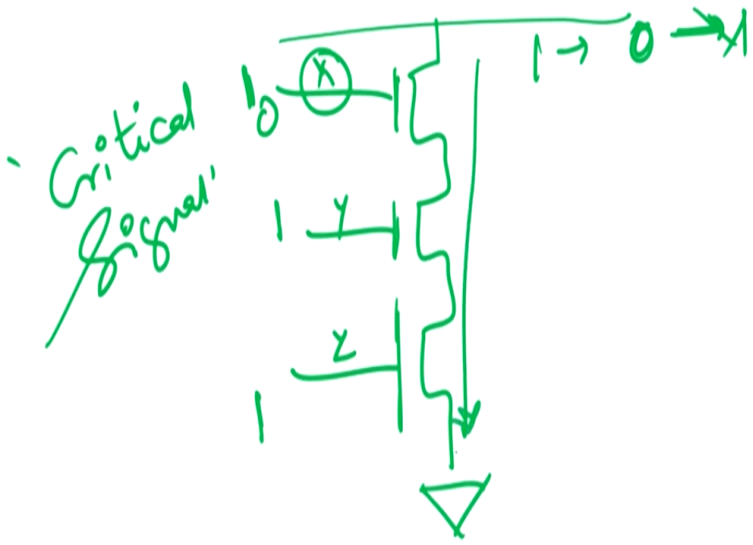
What is high to low? High to low is from this part to this part so what you see is almost quadratic increasing the value of tpHL and the reason is in front of you because for pull down to get activated all the transistor series has to switched on which means that your CL loading increases your resistance of the path increases and therefore tpHL increases right so fan-in as got the inverse or rather than the inverse effect of the delay right. And this you should keep in your mind while designing in any logics right with basic knowledge now available to you or explainable to you.

(**Refer Slide Time: 19:35**)



Let us look at very important issue which is basically your input reordering right if you are looking at input reordering. Now what you do I mean by input reordering I mean to say that.

(Refer Slide Time: 19:46)



I just use to one line first of an we can move forward or I can just see what we have basic issues and we can discuss the other issues later on if I use NAND 2 logic or NOR 2 logic NAND3 logic NOR NAND 2 logic. We define two important properties of a signal the first properties of a signal which you forget about first and second do not worry these two just look at this second one first of all. The signal right last to all the inputs which i have a stable value can be called as a critical signal right which means that signal right which finally say signal coming right and say let me give you an example.

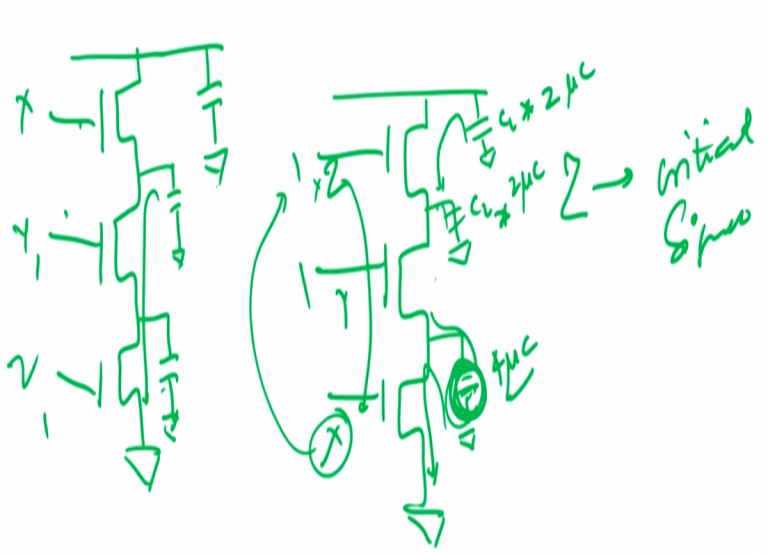
Let us say I have got this right this into consideration I got NMOS in the pull down condition I have NMOS's I have this is one right and this is the signal right x, y, z. x is a signal which is arriving the last as compared to all the three so we defined this to be as the critical signal. So how do you define the critical signal? A critical signal is a signal which arrives the last and based on which the logic is based so if is one the output will go to 0.

If is the 0 the output will go to 1 am i clear so we define the critical signal to be that signal like so critical signal will be defined as just let me erase it so the signal last to all the inputs the last one come to the input which has the stable value can be called as a critical signal on the path and therefore on which the ultimate speed of the critical path which means that and if you go back to

the same definition one appearing here right will not give any output result. But suddenly when this one appear here this 0 sorry this 1 goes to 0 because your pull down path is got activated.

So you have to be very cautious that I define critical signal to be that signal which is basically the last signal to arrive that depending on that I will have my output designated the output value designated. Now rule of term is as I will discuss today the rule of term is generally try to keep the most critical signal closes to the output as much as possible so the rule of term is try to keep the critical signal closest to the output as possible which means that if I have a let us say a 3 input NOR or 3 input NAND.

(Refer Slide Time: 22:35)



Let us suppose and I got 3 inputs here and this is x let suppose y and z then try to keep this z as the critical signal path. And we will discuss why so let us suppose this C1 this is C2 and let us suppose this is CL right. Now you see if let us suppose I kept the most critical signal at x right then what will do so let us come this become 1 when this 2 as become 1 this CL have been discharge to this point this CL will be together discharge and everything will sit here agreed will sit on this last capacitance here.

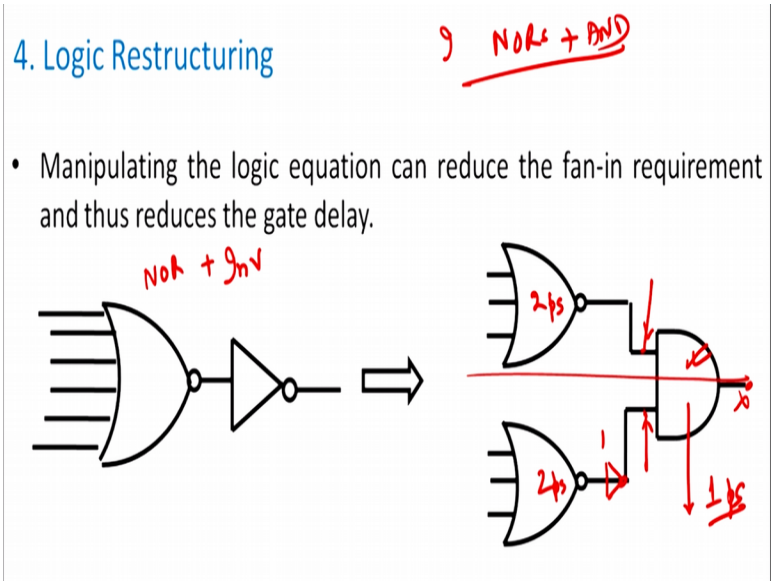
And it has to wait till the signal at x does not become 1 as 0 if it become 0 output this will be equal to 1 if it 1 output will be equals to 0 and then you will have a path available here where as if this would have been the critical signal which is here and this would have been the lower one then by the time so the amount of charge is larger here right why because this has been added to

this so this is if this is the 2 micro column this also the 2 micro column and this is also some so this will be 4 micro column available here.

Whereas I would kept the critical signal right away from my network I would have kept it x, y and z and I would have kept here then if x and y are the both equal y and z are both equals to 1 any cap loading here capacitor loading here will all go to ground right so there is no charge which charge will left the CL charger only left so much smaller charge means much faster dissipation or a much faster you will have much faster drop to the ground. So I thing I made myself clear that the critical signal as to be made as closes to the output as possible right.

And this is very important understanding or explaining the whole issue that makes have a input reordering issues takes care of right another important issues or important idea people have been doing it basically logical restructuring.

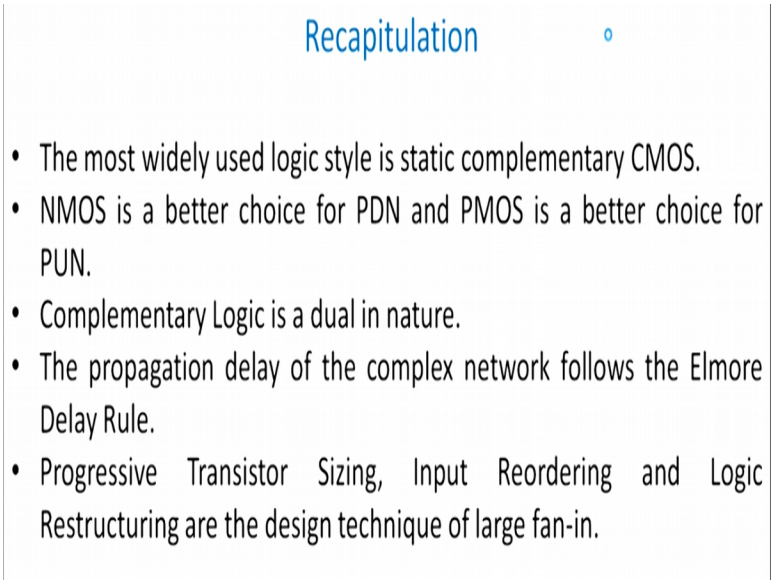
(Refer Slide Time: 25:06)



A logical restructuring is try to make the path as much symmetrical as possible I explain to you what you mean that say if you made a symmetrical then if the delay have two Pico second right then both will arrive here exactly at same instant of time at this point as well as this point right and therefore this will able to evaluate the data without waiting for somebody else where as if there would have been asymmetric for example in this case there would have been inverter also with 1 Pico of second then this gate as to wait at least 1 Pico second of second extra in order to evaluate the voltage or the logic function of the particular point.

So the idea is that either you manipulate the logic function of basically NOR + an inverter right and this basically 2 inverters sorry two NORS + 1 AND gate which have been using it is always advisable asymmetric function in this case this is known as logic restricting. What do you mean by logic restricting therefore logic restructuring you need to restructure your logic so that you symmetric output available to you in this case.

(Refer Slide Time: 26:23)



So let us recapitulate What we have understood this case we started CMOS inverter in part 1, part 2 and we understood what the CMOS logic looks like what are various functionalities of CMOS logic what is the meaning of PUN and PDN pull up network and pull down network and we saw that NMOS is a better choice for pull down network and PMOS is a better choice for pull up network to get rail to rail swing which is VDD is 0 swing available to me.

Complementary logic are dual in nature which means that if the pull down path is on pull up path will be off and vice-verse so both the path cannot be on in the same instant of time and therefore your static power dissipation are minimized to a much lower value and therefore power dissipation are low the price you pay for it is of course your area is almost double because for n input logic you required 2 n number of transistor in such a case right.

We have also seen that the complex network what is known as the Elmore delay rule we have used Elmore delay rule which have just I show to you when I discussing with you NAND 4 logic

calculation of tpHL right and we used the NAND 4 logic do that we also solve that progressive transistor sizing is always nice try to keep the transistor for this away from the output to a larger value.

Similarly we also saw that the critical signal is that signal which is basically the last signal to come before the logic gate settled keep that signal closes to the output so when your trying to size your transistor keep the larger size transistor away from the output and also try to keep the critical signal closes to the output these two we have also understood we have also understood the various design technique of large fan-in and logic restructuring in this case so this takes care of the third module for combination logic law. When you meet next time we will actually starting with the pass logic with the VLSI design right and there will take care of almost come in the all combination logic block thank you very much.