Digital VLSI Testing
Prof. Santanu Chattopadhyay
Department of Electronics and EC Engineering
Indian Institute of Technology, Kharagpur

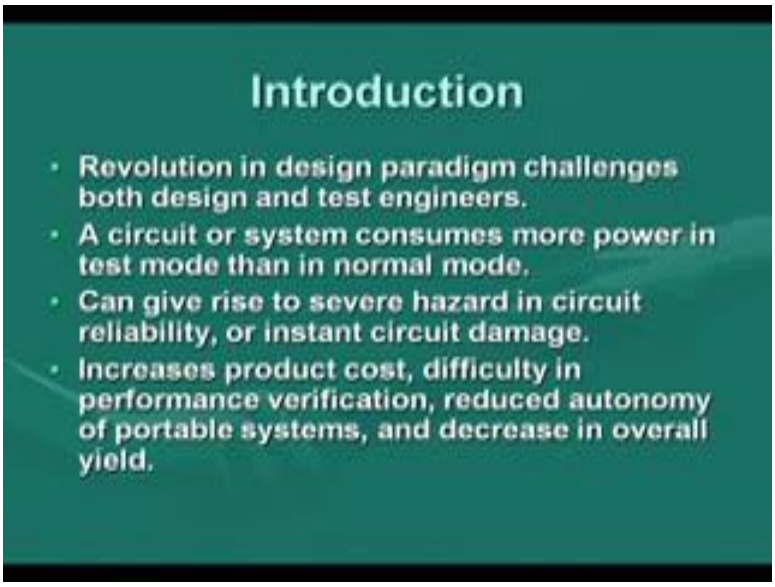
Lecture - 32
Low Power Testing

So, far we have seen different testing techniques for circuits and circuits, so that the objective function that we had was to ensure high fault coverage then sometimes we wanted to generate minimum set of test pattern, minimal set of test pattern. So, that the storage requirement is less then and the fault coverage is high and the test time is minimized. So, in that sense that the time that we have in applying the test pattern that is minimized and for the techniques we have seen that this ATPG tools, they can generate compact pattern sets so that is one process. Then we have seen these compression mechanisms by which we can compact the test patterns. So, we can compress the test patterns and also compress the responses and that way we can reduce the amount of information that needs to be stored into the system.

Now, the point is that with the as the technology is advancing. So, we are putting more and more devices into the system and as a result the other issues are becoming more important than say your test length or test time duration. So, the other issues that are becoming more important are the power consumptions that occur during testing and also the type of temperature rise that occurs during the testing, during testing of the system. Now what happens is that in case of design part of it, design engineers they are bothered about power, they are bothered about temperature, but they have got many tools in their hand. So, they can do something so that some part of the circuit is made to kept of, they also ensured that the several parts of the circuit they do not operate simultaneously. So, that way they can take care of this power problem; however, when we are having this testing. So, it is not always the case the test engineer may still locate these parts should not be done together or say this part should have a separate power supply or separate frequency of operation. So, that may not be possible from the design angle.

So, with this increase in the complexity of testing, power has also become an important issue. So, in this part of the lecture, we will try to see how this low power can be ensured during the test operation.

(Refer Slide Time: 02:51)



So, revolution in design paradigm challenges both design and test engineers. So, revolution has occurred because now what has happened is that complexity of the design has gone up significantly then to cope with this design turnaround time, this reduced design turnaround time, this core based design or this intellectual core based design that has become more popular. So, when it is, for designing a system it is not that we design the whole system together. So, we take the previously design codes from different vendors and try to integrate them into the system. Now if I am designing the whole system then I know which part are important and which functionality of that module is important and I can design my test pattern set accordingly. And we have also seen several schemes like we say that the same test pattern maybe applied to a number of circuits to reduce the testing time in a broadcast scan mode.

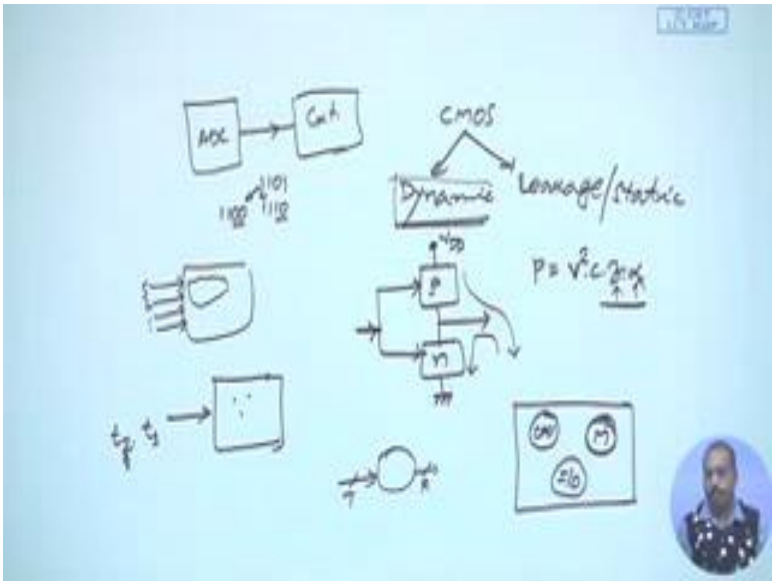
Now, when I am taking codes from different vendor the difficulty that I have is the test patterns can are also to be provided by the vendors because I do not know the internal details of those cores. So, I cannot have my own test patterns set. So, these test patterns

are generated by the core vendors and the core vendors since they do not know in which environment this core will be placed, so they cannot take help of this say this broadcast scan and all those mechanisms that we have discussed in the process. So, they cannot take help of that, as a result they try to give as elaborate a test sequence as possible so that these test. So, that all the faults in the core are going to be tested. So, this creates a large number of test patterns for the whole entire system that we have.

Now, as if test is system integrator, we do not have any other option. So, what we have to do is that we have to apply the test patterns provided by the vendors blindly one after the other and see the responses and compare. So, this makes the whole process very complex and as we have said that testing process anyway occupy about 60 percent of this design time, now with this type of complexity is coming in through the design time has may go down due to this reuse philosophy this testing time will go up further. So, that makes that the operation of testing engineers more challenging. Second observation that we have is that a circuit or system consumes more power in test mode than in normal mode. So, apparently it seems a bit absurd, but in reality that is true.

Because suppose I have got a system that monitors the temperature of a room. So, it senses the temperature values and accordingly depending upon the temperature value it will be controlling some activities, it will be turning on some SCs or turning on heater like that, so it will do something. So, the corresponding circuit, you see it is getting input from some analog to digital converter which is doing this temperature conversion. Now how does this value changes? Like it is not that if you look over a sequence of time, it is not that the value of a tem of a temperature of a room changes abruptly. So, it changes in a rather slow fashion. So, if it changes in a rather slow fashion then for the ADC. So, if I have this is the ADC and this is feeding my system. So, this is the circuit for temperature control now this ADC outputs, they are also correlated.

(Refer Slide Time: 06:30)



So, if the value was say 1101 at this point of time then at the next point of time, it may increase by say 1 bit. So, it may become 1110 or it may decrease by 1 bit becoming 1100 this both of these 2 values that we have, they are related. So, you see the only these 2 bits they are going to change, but the most significant 2 bits they are not going to change.

Now, this is the circuit where these 4 bits are coming as input. So, in this circuit I have got this four bits coming from the ADC given as input and out of this if these 2 bits do not change their value, only these 2 bits change then a large part of the bits that we have in this circuit they will not switch and we know that in CMOS technology the power is consumed only when this switching occurs. In fact, this is due to this dynamic power, in CMOS we have got two types of power consumption; one is known as dynamic power and the other one is known as this leakage power or static power.

Leakage or static power, this dynamic power consumption is coming when the output switches like any CMOS circuit, it can be considered as a p network and a n network; the input is fed to both of them output is taken from the middle. So, n network this is grounded and this is the V DD supply. So, this is the static CMOS type of design.

Now you see that when this output changes to 1 then the capacitor gets charged like this, the load capacitor gets charged like this and when this output become 0 then this load capacitor is discharged like this. But when the input is steady the either p network or n network is cut off as a result there is no current flow through this device. So, that way this current is consumed, power is consumed only when there is a switching that is why in dynamic CMOS, in CMOS the majority of power consumption is due to this dynamic power and this power consumption is actually when the switching occurs. So, when the switching occurs if there is no switching at the input the naturally the circuit gets will not switch as a result they will not consume any power. So, this is what is controlled here. So, we have got, in this particular example that we have taken if this 2 inputs are not changing at all then it is very much likely that this will not be, the power consumption will not be that high.

So, that is reason that is the reason that in case of normal operation of a system the inputs are going to be correlated to each other, say do they do not change much as a result they do not consume, they may not consume much of power. However, when we are putting the system into test mode, in the testing what is required is that if one test pattern is applied, suppose going back to the same example. So, this is the circuit that I have I apply the first test pattern t 1. So, it detects say these three faults. Now the second test pattern if it is similar to t 1 then what will happen. So, it will possibly detect say these three faults and maybe some more fault or maybe some less fault that way it will work, but it is very much unlikely that it will detect a drastically different set of faults if the second pattern t 1 is very much similar to t 2 is very much similar to t 1.

So, as a test engineer what we will like to do is that t 1, t 2 it should be highly different from t 1. So, t 1, t 2 should be highly different from t 1, so that it excites some other faults which were not excited by t 1. So, that is the exception. So, in case of testing what is happening is that this successive patterns that we are having, successive patterns are highly uncorrelated whereas, for this normal operation successive inputs are often very much correlated to each other. So, that is why this circuit or system so that consumes more power in test mode than in normal modes. So, if the correlation is less; that means, there will be more or number of switching that will occur because many of the inputs will change as a result it will consume more power.

So, it can give rise to severe hazard in the circuit reliability. So, reliability may be a question because of this power consumption, a good amount of power is converted to heat and that way it will increase the temperature and if the temperature increases then the delay of the circuit will also change.

So, if a circuit is not uniformly heated or portions of circuit are not uniformly heated then what is going? So, the delays of different parts maybe different, when we are doing a delay test it may so happen that a chip will fail in delay test and though the chip is actually correct, but due to this heating this delay difference is have occurred and as a result the chip gets discarded. Or it may so happen that some chip which has got this delay problem, but due to this uneven hitting it passes through the delay test and a good chip a bad chip is being shipped to the customer, so that can happen. Or it can cause instant circuit damage, if the power, if the power f d circuit has got a upper limit, if the power limit is exceeded then it will cause a cause a damage to the circuit.

Increases product cost because now I have to beep more causes in my testing process. So, then difficulty in performance verification and then this reduced autonomy of portable system and decrease in overall yield, this power ever testing if we want to do then now we have to test some special care. So, it cannot test maybe we wanted to test a number of components together to speed up the testing process, but it may so happen that the combined power consumption is crossing the power limit as a result we want to restrict that. So, we cannot do that parallel testing. So, that way the testing time will increase. So, that will affect the product cost.

Performance verification is in problem because again the same thing that is this power ever operation has to come into picture portability of; portability will reduce because now the circuit has to be the battery operated circuit if it has to be tested periodically and the power consumption is high for the test power consumption is high then this battery life is in question. So, this autonomy of portable system so that becomes less and as a result the overall there will be a decrease in the yield.

(Refer Slide Time: 14:00)

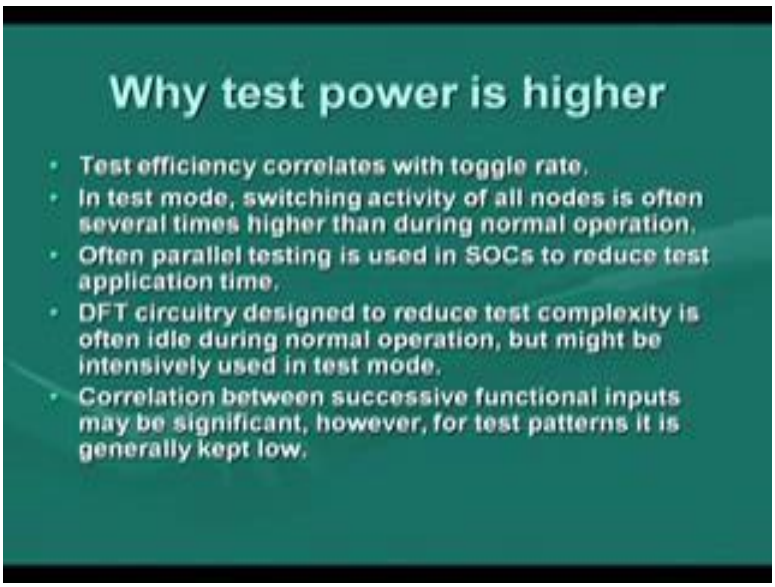
So, what are the challenges that test engineers are facing due to this changes like escalating transistor count increasing chip complexity while maintaining its size. So, this is always there because where design is becoming more and more complex. So, this issue will come into picture. Transistor count will increase, then the complexity of the chip will increase and the chip area is remaining unaltered. So, as a result the density or devices part unit area so that is going to increase, so that is going to increase the testing complexity.

And also testing is one of the most expensive and problematic aspects in a circuit design cycle. So, as we have seen, we have discussed previously that it is taking about 60 percent of the time and nowadays what has to be done is as testing goes in parallel with design so that both of them go simultaneously, previously it was done at the end after the design is over then the testing process used to start, but now it is not like that, but this is one of the very problematic aspect.

So, whatever we have techniques we have seen so far that is the traditional techniques, there the test methods are evaluated based on area requirement that is if we are proposing some DFT what is the extra area that is required, so if you are proposing some test pattern generation algorithm then what is the fault coverage that it can produce. Then

what is the test application time, so how much time we have to give for applying the test pattern, this test pattern set. So, these are the factors which were determining the quality of a testing scheme or including the architecture and the test pattern set. But these low power systems make power management a critical parameter because low power systems they will have a power budget, so you cannot go beyond certain power consumption. So, normally the designer will tell the power budget for a system. So, it cannot go beyond that. So, as a test engineer, we cannot apply patterns so that the power for that budget is exceeded. So, that way this power management becomes an important part in the testing process also.

(Refer Slide Time: 16:18)



Why test power is high? First one is the test efficiency correlates with toggle rates. So, this point we have discussed a bit. So, what it says is that if I apply fast test pattern that excites some faults the next test pattern that I would like to apply that should excite some other faults otherwise it will not detect new faults. So, this pattern that we are going to use must be a totally a different one from the existing ones. So, that way this test efficiency it correlates with the toggle rates. So, that is one very important thing.

Then in test mode switching activity of all nodes is often several times higher than during normal operation. So, this switching activity is actually a measure like even

though the system is operating so how many times the output changes. As we have seen that for CMOS circuit these RT analyst this output changes this dynamic power consumption is not there and till say 90 nanometer technologies, we have got 80 percent of power consumption of circuit is due to this dynamic, due to this dynamic power and rest are due to static power. Of course, with expansion with the advancement in technologies the situation is changing. So, now, this leakage power is also a very important part of this whole process and the problem with leakage power is that for about leakage power consumption. So, that way part of it will be converted into heat and that heat a for every 10 degree centigrade rise in temperature this leakage is going to almost double itself. So, that becomes say positive feedback on the system so that will be catastrophe. So, that will call, that may cause a thermal runaway.

So, we will come to those issues later, but this testing mode this power consumption minimization becomes an important issue. So, the switching activity of nodes are several times higher because of this requirement that is in case of normal operation the successive inputs are correlated to each other, so many inputs do not change for the module whereas, for this testing part this successive test patterns they are highly different from each other as a result they will cause this more more number of toggles.

Often parallel testing is used in SOCs to reduce test application time. So, what do we mean is that in a system on chip design what has been done is that if this is your chip that you are manufacturing then in this chip we have got different modules from different vendors maybe we have got a CPU core from some vendor, a memory core from some vendor then one IO core from some vendor. So, these are the different modules that we have.

Now, you see that if since this modules are not designed by me, so I do not know their details. So, I cannot generate test patterns for them. So, the test patterns have also been provided by the core vendors. So, core vendor has said that if this is the CPU. So, if you apply these set of test pattern T and if you get this response R; that means, the system the CPU is correct, but we cannot take this description as it is because if it is a board level design then the advantage is that you take the individual chips from the vendors and the vendors they have ensured that the chips are correct. So, what you meet to test is only the

interconnects between them, but in this particular situation what is happening is that we have got only a soft description of the CPU in terms of netlist from the vendor and that we have combining in to the one on a silicon floor we are going to manufacture them.

Now even if my manufacturing process is 100 percent correct, so they are maybe impurities in the silicon, no silicon is 100 percent pure so there may be impurities as a result the CPU might not have been manufactured, might not have been fabricated properly.

So, what is required is that I need to test the CPU even if my manufacturing process is 100 percent correct. So, that way this testing is required. So, what we do? Since each vendor has given me a large set of test patterns and if I test this individual cores one after the other so it will take lot of time. So, we try to do a parallel testing. So, with parallelly test a number of modules number of cores together, but the as a result what will happen is that in normal operation maybe those cores are not going to act together, they are not going to act in parallel. So, even if they are high power consuming cores the designer will know that these cores are not going to be activated simultaneously as a result the power management policy will switch off power for one of them. So, power will be, the system power will be within the limit.

But, but as a test engineers. So, what we are trying to do here? We are trying to test a number of cores together parallely, when we are doing that then we are going to have the power consumption of all these modules and if 2 high power consuming cores are tested in parallel then we have got this power consumption, there is a possibility that the total power consumption will cross the power budget of the system. So, this is the point that often parallel testing is used in SOCs system on chip based designs to reduce the test application times. So, that will increase the power consumption.

Then this DFT circuitry design to reduce test complexity is often idle during normal operation, but might be intensively used in test mode. So, that we have already seen that the scan chain and all that, they are actually not active when the circuit is operating in the normal mode. So, then this is only the normal operation are going place. So, the there are some error detection circuitry maybe there then the beast that requires this test pattern

generated and response analyzer to work. So, those modules are not active in the normal mode of operation of the circuit.

Whereas for test mode of operation. So, they are coming into picture. So, in this test mode of operation this circuitry that will be activated so that will reduce the test complexity, but they will require, they will be intensively used during testing if so the power consumption of those circuits will cause the overall system power consumption. Correlation between successive functional inputs maybe significant; however, for test pattern it is generally capable. So, this point we have already discussed. So, successive patterns may be input patterns maybe correlated, but successive test patterns they are often highly uncorrelated.

(Refer Slide Time: 23:22)

So, what are the solutions? There are some ad hoc solutions; ad hoc solutions are like over sizing the power supply. So, increase the power supply size because we need for testing we need higher power. So, maybe we have got a separate power supply for testing. So, that is one possibility.

Then the packaging, packaging is an issue because I have to have some packaging that will that will be able to sustain the extra heat that is produced so that way the packaging

maybe more bit, packaging maybe made better increasing the cost of the system and cooling maybe there, so we may have some cooling facility. So, that it can withstand the increased current that is coming in the test mode. But each of them have got their problem, like nobody will like to have a large sized battery for providing the power supply. Then for testing purpose nobody is going to give us a packaging facility is costly and cooling is of course, another problem like. So, we cannot say that there will be a fan fitted because this chip will become hot during testing. So, for during test we want to cool it, so we will have some fan. So, that is not possible. So, this is one type of solution.

Another solution is to insert breaks in the test process to avoid hot spots. So, if the circuit is tested continually and if one region of the circuit is excited throughout the test session then what will happen is that it will that part will become hot. So, that creates some hot spots in the circuit. So, what we do? We break the test session, we take, maybe if total 1000 patterns are to be applied maybe we apply 100 patterns take a break then again we apply 100 patterns take some break, so that will allow the system to cool down. So, when it cools down then this static power; leakage power consumption will come down. So, as a result this testing part test power can be reduced. Then testing with reduced operating frequency, this is the other possibility like you see you know that the dynamic power consumption that we have in a CMOS circuit. So, this is proportional to v square the supply voltage into c into f into alpha where v is the supply voltage c is the capacitance that switches f is the frequency of operation and alpha is the switching activity.

So, if we can restrict these values. So, out of these supply voltage we normally cannot do anything because that is technology dependent, it is already done. Similarly c part that is the capacitance that switches, so that is also depicted by the design because designer has put the output of one gate going to 10 different places. So, all those 10 gates input capacitance will be seen as the load capacitance of this particular gate. So, as a result, we have got this c value is also not in our hand.

Then this f the frequency of operation, this is one thing that we can reduce. So, we can say that for testing we will do it at a lower frequency though the actual circuit will operate in the gigahertz range. So, maybe we will be doing our testing in the kilo hertz range or the mega hertz range and this alpha is the switching activity. So this is another

ad hoc solution, we test at a reduced operating frequency, but again the problem is if you do that then all this delay related faults they cannot be captured. So, this delay faults will not come into picture. So, we they will not be detected, so that is our problem.

And we can also do a partitioning of the system. So, we can divide the system into different parts. So, we can test the system one part at a time. So, that is another possibility, but again the thing is that how easy it is to partition a system so that it depends on the system that we have in our hand. So, we need to insert some extra multiplexers and other modules so that this partitioning can be affected. But whether, how easy is doing this partitioning, so that remains a question. And test planning, test planning is very important like when will we test which module, for how much time, etcetera, this has to planned properly. So, we continue in the next class.