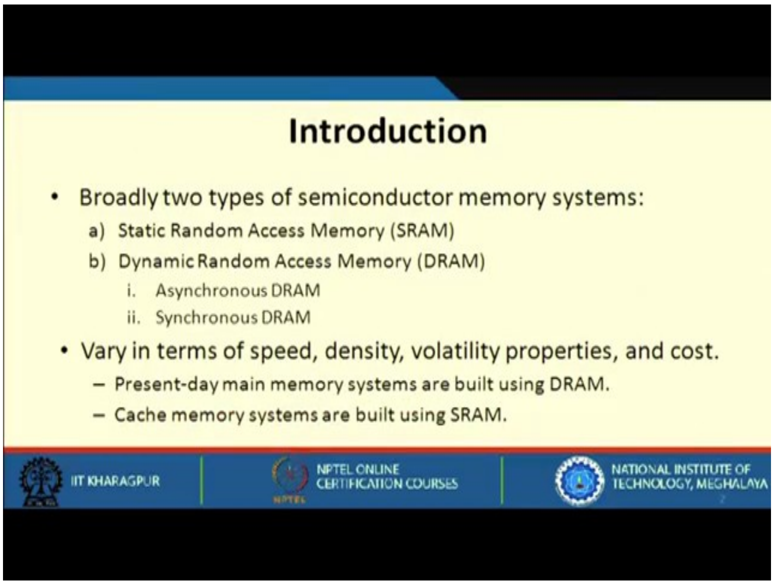
Computer Architecture and Organization
Prof. Kamalika Datta
Department of Computer Science and Engineering
National Institute of Technology, Meghalaya

Lecture - 24
Static And Dynamic Ram

Welcome to lecture 24. In this lecture we will be looking into static and dynamic RAM. So, broadly two types of semiconductor memory systems will be seen. And we will be seeing how a single-bit SRAM or DRAM cell is built. How a single bit is built and then how you can extend it to any size?

(Refer Slide Time: 00:55)



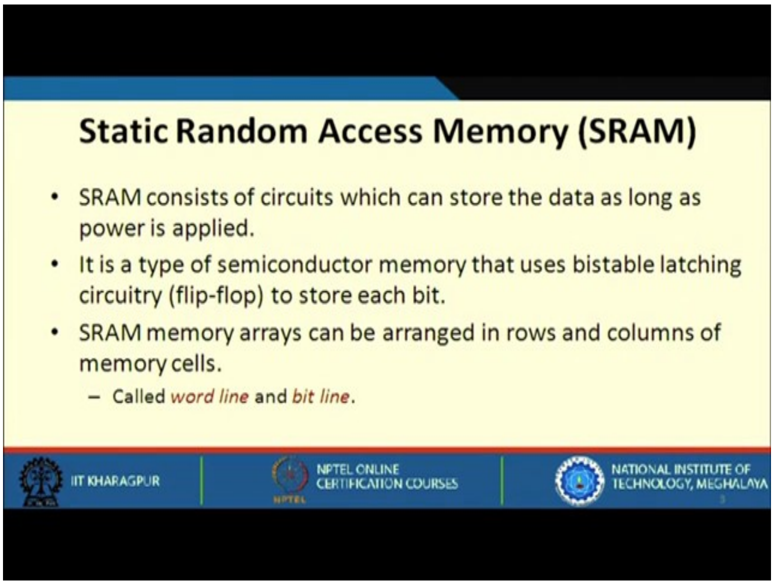
So, as I said broadly two types of semiconductor memory systems exist, static random access memory and dynamic random access memory. Under dynamic random access memory, we have two more types; one is called asynchronous DRAM, another is called synchronous DRAM.

Now, how these how these types of semiconductor memory vary? How you can differentiate among them? They vary in terms of speed; that means, how fast it is? How fast is your SRAM? Or how fast is your DRAM? In terms of density; that means, within the same area, how many bits you can pack using SRAM? Or how much bits you can pack using DRAM? That is meant by density. Then comes volatility.

So, if you have power supply given, then static RAM will have its data. But in case of dynamic RAM, even if you have supplied the power the data may not retain. You have to do periodic refresh to keep the data. So, these memories vary in terms of speed, access time, density, volatility, and cost. We will see that all these things are very much related to each other.

We will see the pros and cons of these properties. In present day our main memory system are built using DRAM, and cache memory systems are built using SRAM. And we will see that our cache memory system is relatively faster, but it is small. And our DRAM is relatively slower,

(Refer Slide Time: 03:53)

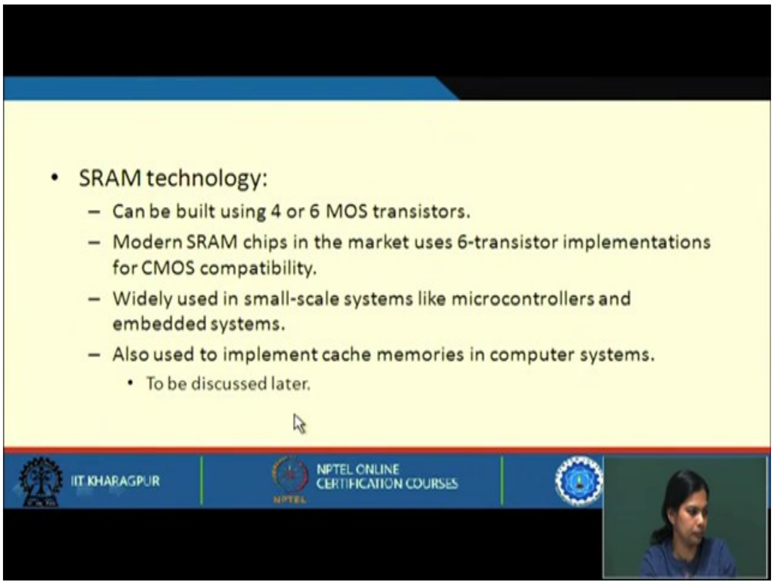


but it can store a large amount of data. Coming to SRAM, it consists of circuits that can store the data as long as power is applied. In this type a semiconductor memory, a bistable latch circuit or flip flop, is used to store each bit of data.

To store individual bits of data bistable latching circuitry is used. And SRAM memory arrays can be arranged in rows and columns. We have already seen that how a memory system is organized. A memory chip is organized in terms of rows and columns.

So, SRAM memory arrays can be arranged in rows and columns. And these are called word line and bit line.

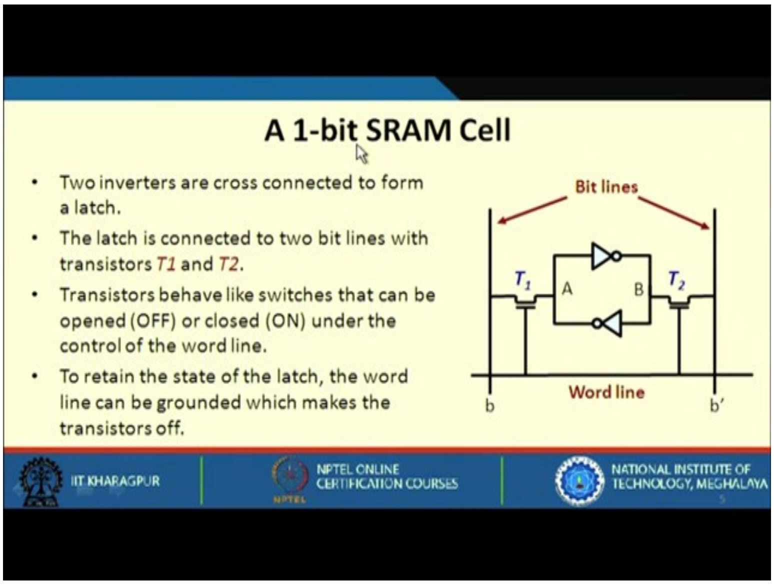
(Refer Slide Time: 05:27)



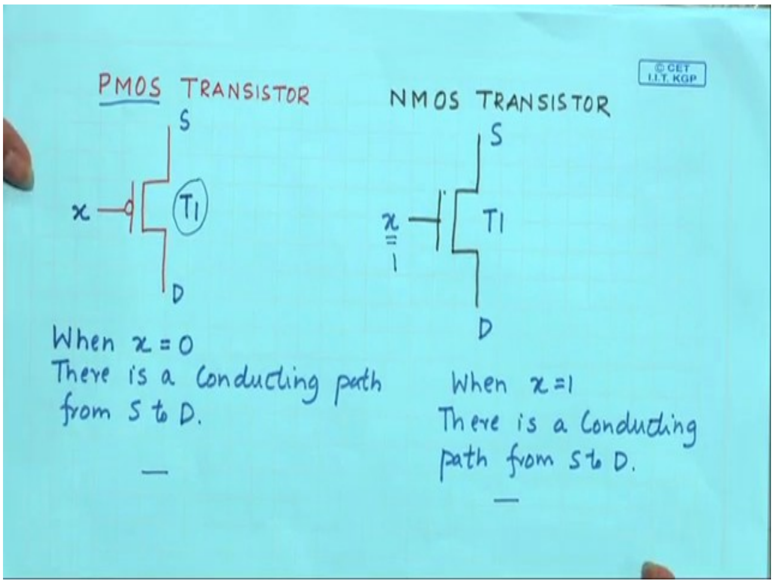
Now, SRAM memory cell can be built using 4 or 6 MOS transistors. But modern SRAM chips in the market it uses 6-transistor implementation or CMOS compatibility. And this kind of SRAM chip which uses 6-transistor are widely used in small scale systems like microcontrollers and embedded system. We know that today's microcontrollers and embedded systems are everywhere.

It does not require very large memory, but it requires faster memory. So, SRAM chips are used in these microcontrollers and embedded systems. And are also used to implement cache memory in the computer system which will be discussed later.

(Refer Slide Time: 06:33)



(Refer Slide Time: 06:39)



Now, before going to this slide I will just discuss two things, one is PMOS transistor, another is NMOS transistor. In this PMOS transistor you can see that, there is an input that is x, and this is source or drain. Basically, this transistor acts as a switch. And this is controlled by this input. If this x input is 0 for PMOS transistor there is a conducting path from S to D.

So, T1 is on. If x is 0 then there is a conducting path from S to D and this will make T1 on. This is the feature of PMOS transistor. Now see the NMOS transistor. In NMOS transistor in the same way, if this value is 1 then only there is a conducting path from S to D. So, what is the difference between these two? In PMOS if the input is 0, then only it will be conducting. And in NMOS if this input is 1 then only this transistor will be on. That is, there will be a conducting path from S to D.

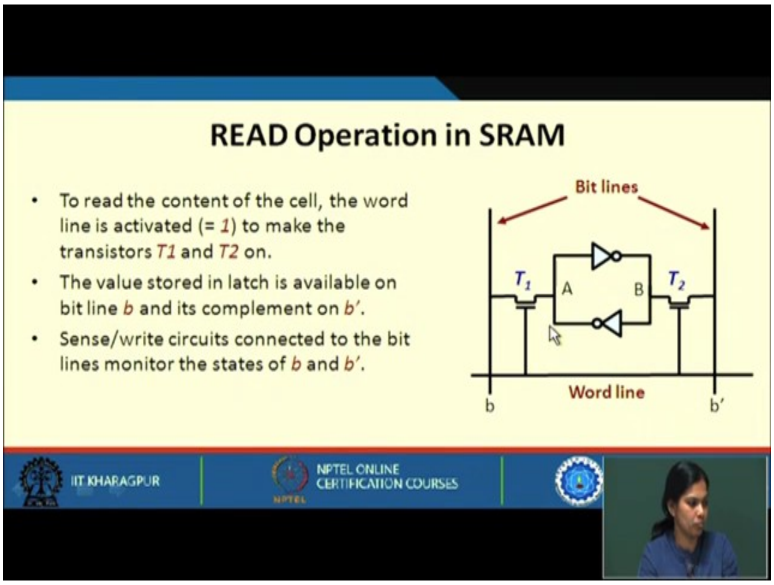
Now, come to a 1-bit SRAM cell, how it looks like. This is a single bit SRAM cell. Here we have two inverters; this is a symbol of inverter. So, two inverters are cross connected to form a latch. So, if you give a 0 input here this will be 1. If you give a 1 input here this will be 0. And now this latch is now connected to two transistor T1 and T2. Now what is T1 and T2? You see this is a NMOS transistor.

So, if you want to activate this transistor, what input you have to give? You have to give a 1 input here. Then only this transistor will be on. So, this particular latch is connected to the 2 bit lines. If you recall our discussion, each of the cell is connected to 2 bit lines, b and b-bar. In the same way, this is a single cell which is connected to 2 bit lines b and b-bar, through transistors T1 and T2.

And these transistors behave like switches.

To retain the state of the latch, the word line can be grounded that makes the transistor off. Now if we want to retain the value which is there in the latch, in that case what we have to do? We have to give this word line to ground. If it is ground this will be 0. Then there is no conducting path. Whatever value is in the latch, it will remain there. So, we have bit lines, two inverters are cross connected to form a latch, and these are connected to 2 transistors through to the bit lines through this transistor, and this transistor can be made on and off. Depending on that, we can read the value or we can write the value into the cell. Let us see that.

(Refer Slide Time: 11:59)

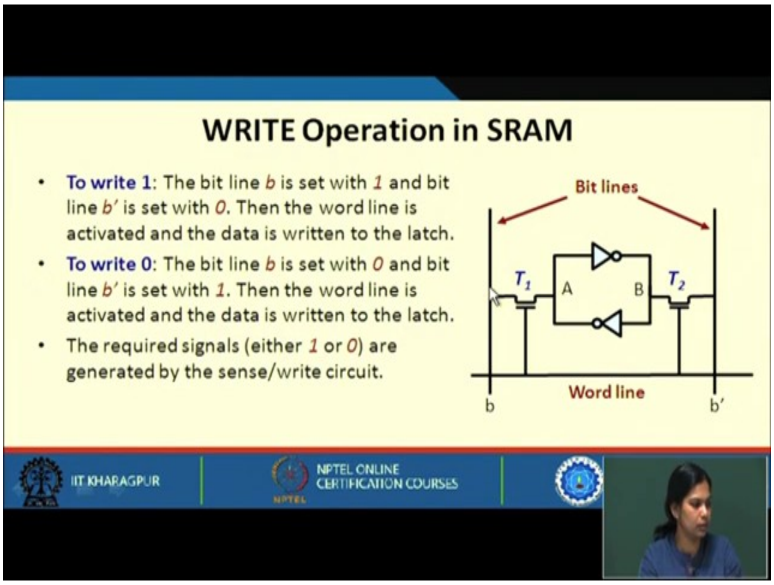


Read operation. How we will read the value? That means, if the value here is let us say 1, then 1 will be at this A, and 0 will be at this B,, because the bit line b will have 1 and b-bar that is the complement will have 0.

To read the content of the cell, what we need to do? The word line is activated. So, now, I want to read this content whatever is in A and B. So, we need to activate this word line. By activating means we are supplying 1 here. So, if we make it on if we activate this word line, then what will happen? This transistor T1 and this transistor T2 will be on. If this transistor T1 and T2 is on, then the value which is stored in A and B that is in the latch will be available on bit line b, and will be available on bit line b-bar.

So, if the value is 1, then 1 will be available in the bit line b, and 0 will be available in bit line bbar. In the same way if the value is 0, then 0 will be available in bit line b and 1 will be available in bit line bbar. And then a sense or write circuit connected to the bit lines will monitor the state of b and bbar. And accordingly it will figure out whether it is 1 or 0. So, this is how we perform read operation here.

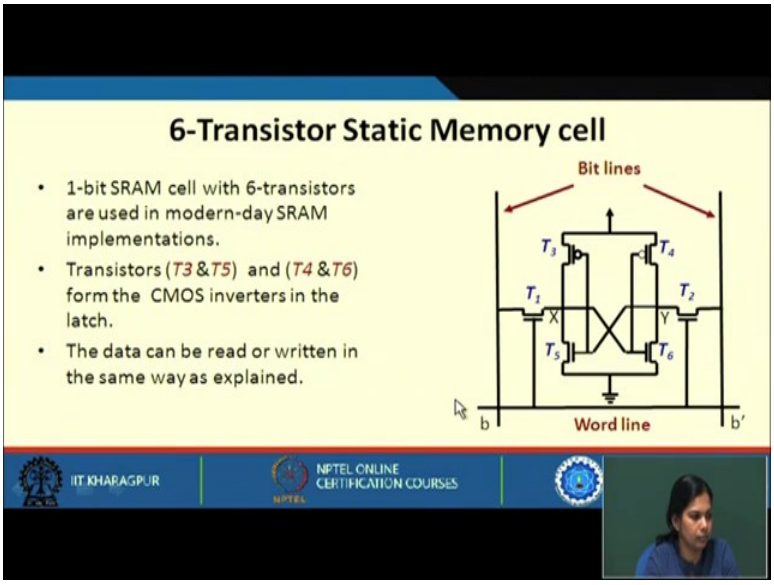
(Refer Slide Time: 14:18)



Now, moving on let us see the riwriteght operation in SRAM. Now for writing we can either write 1 or we can either write 0. First see if I want to write 1, then what I need to do? I need to set 1 in bit line b. In this b I have to set 1, and I have to set 0 in bit line bbar. So, bit line b will have 1, bit line bbar will have 0. Then I will activate the word line. Which will make the transistor T1 and T2 on whatever value is in the bit line will be available in A. And whatever value is in bbar will be available it B. So, the data is written to the latch.

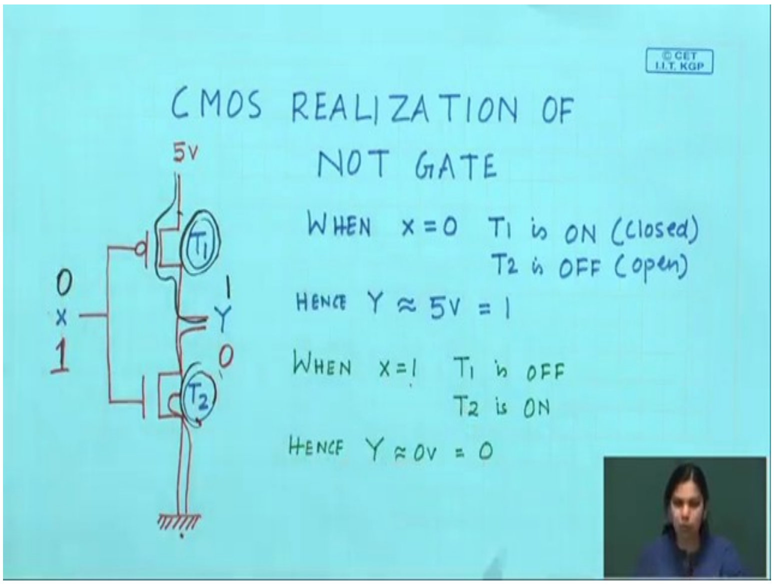
Similarly, if I have to write 0, then I apply 0 in the bit line b, and 1 in bit line bbar. And in the same way I activate the word line. By activating the word line the transistor will be on and whatever data will be in b will be stored in this latch, and whatever will be in the bbar will come here. So, this latch will now have the value which is there in these bit lines will be available in this latch. And now as I said, I can either write 0 or I can write 1; the required signals that is either 0 or 1 will be generated by the sense or write circuit. So, this is how write operation happens in SRAM.

(Refer Slide Time: 16:08)



Now, see before moving here, if you consider this diagram you have a NOT gate. So, let us see the CMOS realization of NOT gate. Here is the CMOS realization of NOT gate.

(Refer Slide Time: 16:32)



You can see here that this is the circuit, or PMOS transistor. A PMOS transistor is connected here. This is T1 and an NMOS transistor is connected here which is T2. Now, this is a NOT gate. Let us see how this will act as a NOT gate; it means if I give x input as 1, the output should be 0. If I give x as 1, then the transistor T1 will not be conducting. Because, this is a PMOS transistor and it will only conduct when the input here is 0.

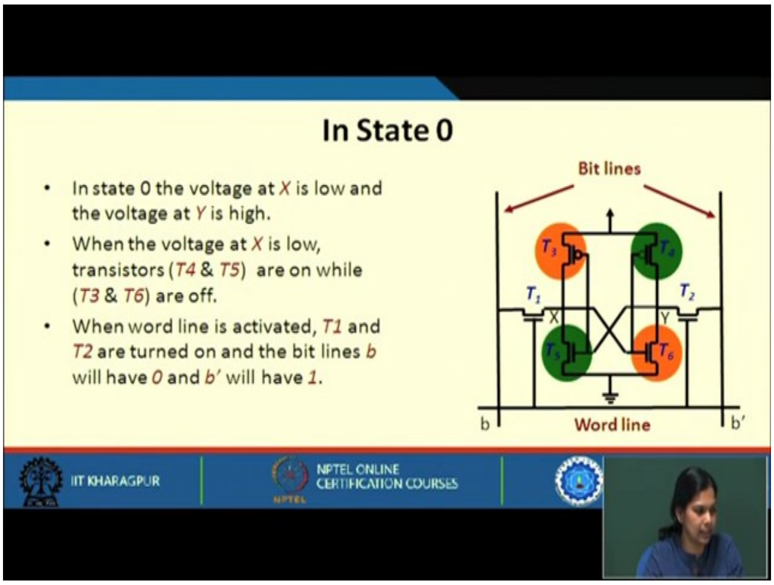
So, there will be no connection from this is not connected. This is open. T1 is off. Now if this is 1, the bottom transistor T2 which is a NMOS transistor will be on. And there is a path from this Y to ground. So, if there is a path from this Y to ground then this Y will have the value approximately equals to 0. So, when this x equals to 1, T1 is off, this transistor is off, and T2 is on. Hence Y will be have approximately a value of 0 volt, which is equivalent to 0.

Now, let us take x as 0. If x is 0, I must get the output as 1. If this is 0, then the above transistor that is T1 will be conducting, but the below transistor that is T2 will not be conducting. If this is conducting there is a path from this 5 volt to Y. So, this Y will have roughly equivalent to 5 volt, which is equivalent to 1. So, if we give input 0, the transistor T1 will be conducting and we will have an output 1. And if x is 1 then, the below transistor will be conducting and we will have the output Y.

Now, what we will do? Once I have shown you this CMOS realization of NOT gate. Now moving on, we will see 6-transistors static RAM cell. This is the CMOS realization of NOT gate. So, I have just replaced this NOT gate with the CMOS realization. And now what we are getting? We are naming these various transistors. So, this is transistor T1 and T2 initially it was there. Now this transistor is T3 and T5. And this transistor is T4 and T6.

Now, one bit SRAM cell with 6-transistors are used in modern day SRAM implementation. So, you can see T3 and T5 and T4 and T6 forms the CMOS inverters that I have just now explained. And reading a data, the data that is to be read or written can be done in the same way as explained in the previous example.

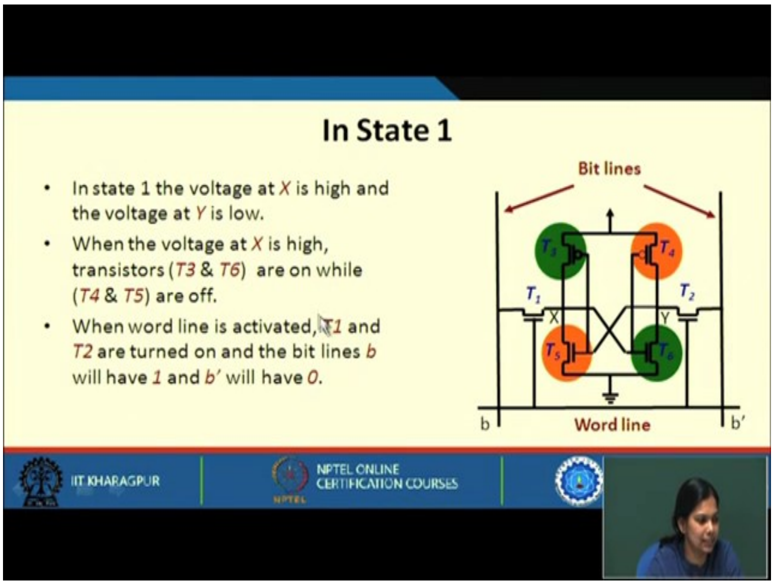
(Refer Slide Time: 21:00)



Now, let us see in state 0 what happens. In state 0; that means, this x is having 0 and this Y is having 1. In state 0 the voltage at X is low. So, here it will be low and the voltage at Y is high, X is low. So, X input is going to this T6 and it is going to this T4.

So, if it is going to this T6 and this is 0, then this will be off. But this T4 will be on. And now Y input is having 1. So, if this is 1 then this T5 will be on, but this T3 will be off. So, that is what T4 and T5 will be on, while T3 and T6 will be off. Now when the word line is activated T1 and T2 are turned on. And the bit line b will have a 0 value.

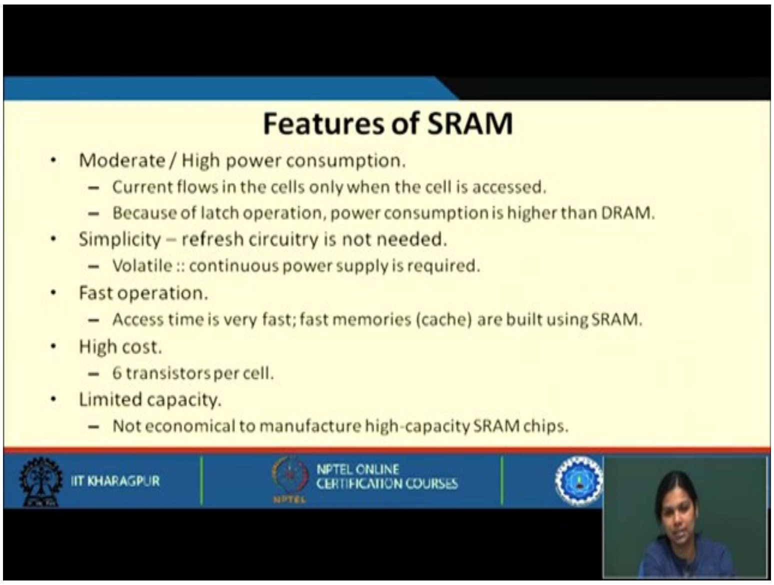
(Refer Slide Time: 22:53)



So, this is how in state 1 it happens. Similarly state 0 it happens like this. Now let us move on what happens in state 1. State 1 means X will now be high and Y will be zero; that means, the bit line b should have 1 and bbar will be 0. So, as X is at 1. So, X is going here and here. So, this will be off, but this will be conducting. And similarly Y is 0. So, Y is low or 0. So, this will make this as non-conducting, but this will be conducting. So, T3 and T6 will be conducting, and T4 and T5 will be off. Now when the word line is activated in the similar fashion T1 and T2 will be on will be turned on. And the bit lines b will have 1 and bit line bbar will have 0.

So, this is how what happens in state 1. Let us see some features of SRAM. So, here the current flows in the cells only when the cell is accessed. This is a CMOS cell property.

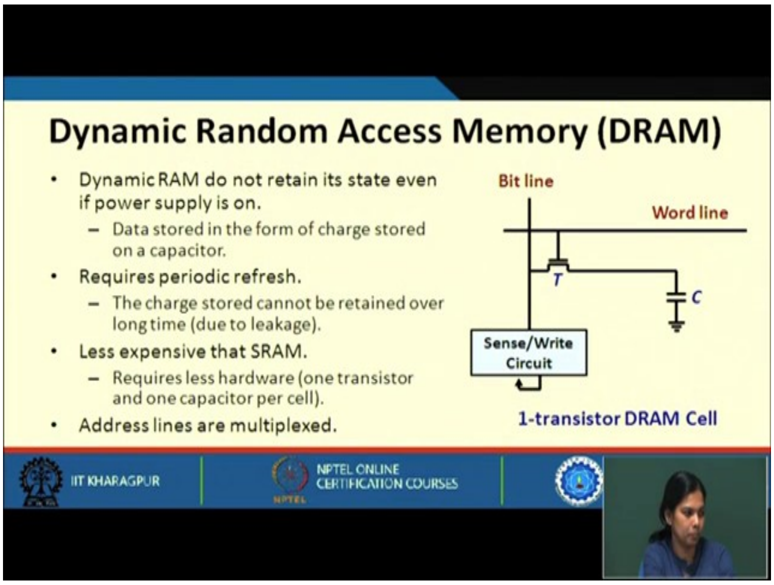
(Refer Slide Time: 24:00)



So, current flows in the cells only when the cell is accessed. Because of this latch operation power consumption is little higher. What is the simplicity? No refresh circuitry is required. It is of course volatile but as long as the power is supplied to it, you need not have to do any kind of refresh. It is much faster. Access time is very fast. So, the fast memories like cache are built using these kinds of cells. But the cost is high. Why? We see that here we require 6 transistors to build 1-bit memory; also the space it takes is more.

So, the cost is high. And of course, it has got limited capacity. Because we cannot built a very large a SRAM cell, as it requires 6-transistors per cell.

(Refer Slide Time: 25:27)

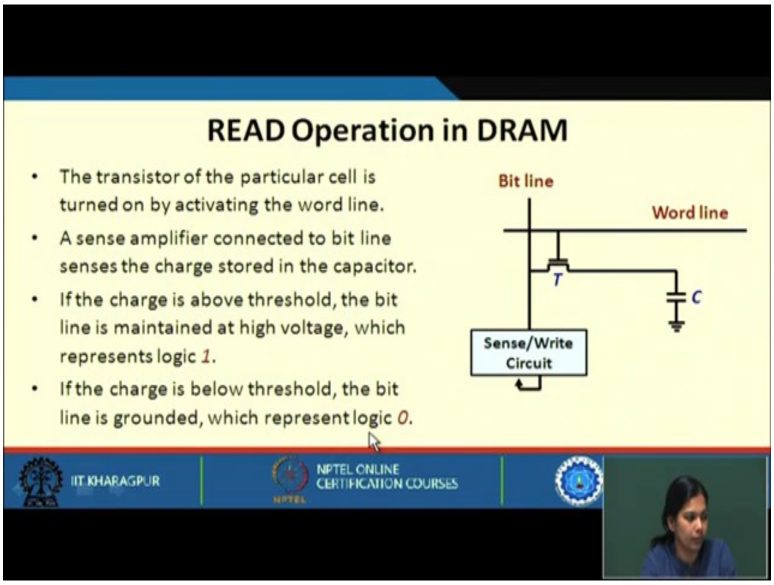


Next coming to dynamic random access memory: in dynamic random access memories as we know that it do not retain the state even if power is supplied to it. So, here the data are stored in the form of charge on the capacitor. And this charge cannot be stored for longer period of time. And this happens due to some leakage property of the capacitor as well as this transistor.

But you see; how simple is the cell. You have one transistor that is connected to the bit line and this transistor is also connected to the word line. Because through word line it will get activated and the transistor is also connected to this capacitor which is grounded. And then it is connected to the sense or write circuit. This is a 1-transistor DRAM cell. But it requires periodic refresh because we are storing the data as charge in the capacitor. And this charge can be retains not for longer period of time.

These are less expensive, we can see that because only one transistor is required to built it; one transistor and one capacitor is required to build this. And here the address lines are multiplexed we will be seeing this little later.

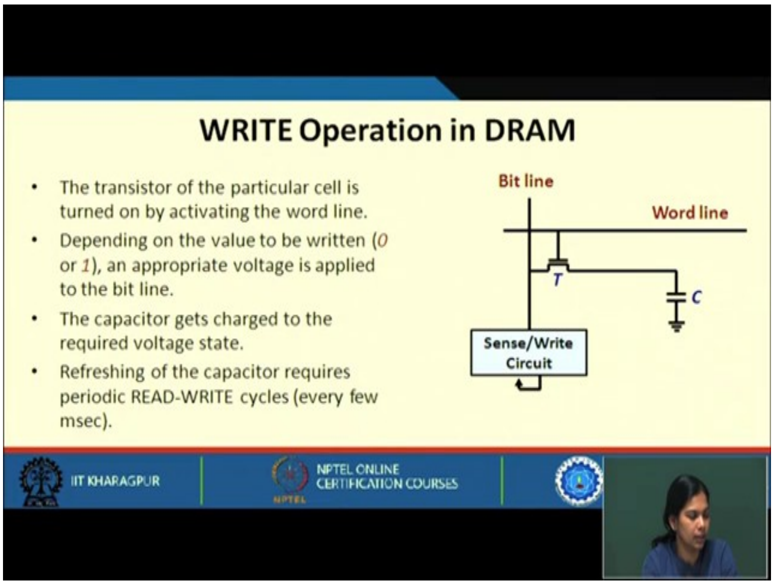
(Refer Slide Time: 27:04)



Now, let us see how we can read the value here. So, as we said that the data is stored as a charge in this capacitor, and that represent whether a 1 is stored or 0 is stored. So, let us see for reading the data from this cell the transistor of this particular cell is turned on by activating the word line. So, this is the word line we activate the word line, such that this particular transistor T is on. Now we have a sense amplifier connected to the bit line. And this line it senses the charge stored in the capacitor. Once this is on there is a connection between this bit lines to this through this transistor.

If the charge is above certain threshold, then we say that the bit is maintained at high voltage, that is 1. And it will represent logic 1. If the charge is below certain threshold, then we say that the then the bit line is grounded, which represent logic 0. Now we see that if we read a cell, automatically it is getting refreshed. Because, we are keeping the required data that is to be stored in this capacitor. So, if it is 1 then this is made on and we sense the charge in the capacitor that automatically refreshes. In the same way if it is 0, then also automatically refreshes.

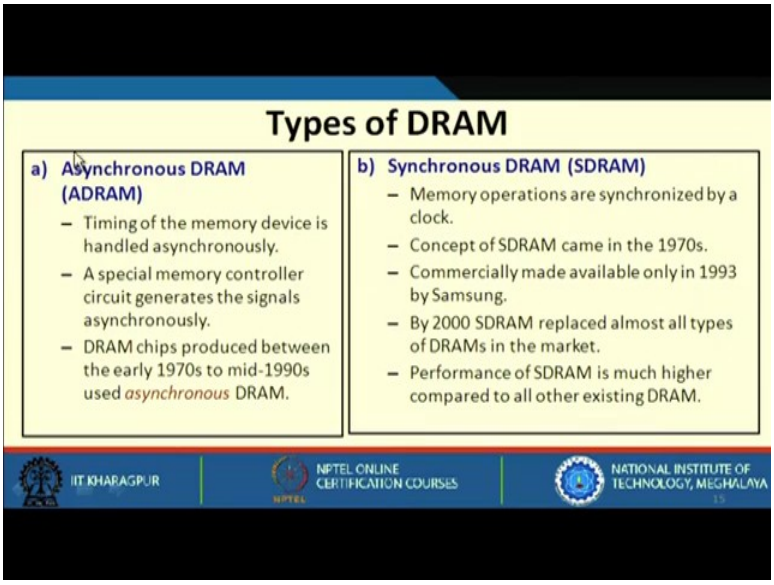
(Refer Slide Time: 29:06)



In the write operation what happens? Through this sense or write circuit this bit line will have the available data, either 0 or 1. The transistor of this particular cell is turned on by activating the word line and depending on the value that is to be written, either you have to write 0 or 1, an appropriate voltage is applied to this bit line.

And, as an appropriate voltage is applied to this bit line, this capacitor gets charged to the required voltage. If it is 0 a required voltage it is charged to that required voltage if it is 1, it is charged to that particular required voltage. And refreshing of the capacitor requires periodic read/write cycles every few milliseconds; so in every few milliseconds if even if you are not reading you have to do refresh to store the data to keep the data.

(Refer Slide Time: 30:07)



Now, there are various kinds of dynamic RAM. One is asynchronous DRAM another is synchronous DRAM. As the name suggest, in asynchronous DRAM the timing of the memory devices handled asynchronously. What do you mean by that? Here there is no fixed timing; I mean when you give a request for read and when the data will be available. The processor has to take care of when the data is available.

But in case of synchronous DRAM it is not like that; there is timing involved to it. And after a particular time data will be available. In asynchronous DRAM a special memory controller circuit generates the signal asynchronously. The DRAM chips that are produced between early 1970s to mid 1990s all used asynchronous DRAM, but today's computers all use synchronous DRAM. So, here the memory operations are synchronized by a clock.

A clock is there which synchronizes it. And this concept was already available in the 70s, but commercially it was available in 1993, and by 2000 SDRAM replaced almost all types of DRAMs in the market. So, there is no asynchronous DRAM these days. We have all synchronous DRAM. And the performance of SDRAM is much higher compared to all other existing DRAMs. So, we have seen in this lecture what all semiconductor technologies are used to build SRAM and DRAM. What the various kinds of DRAMs? And now we will be seeing specifically the various kinds of DRAMs.

Thank you.