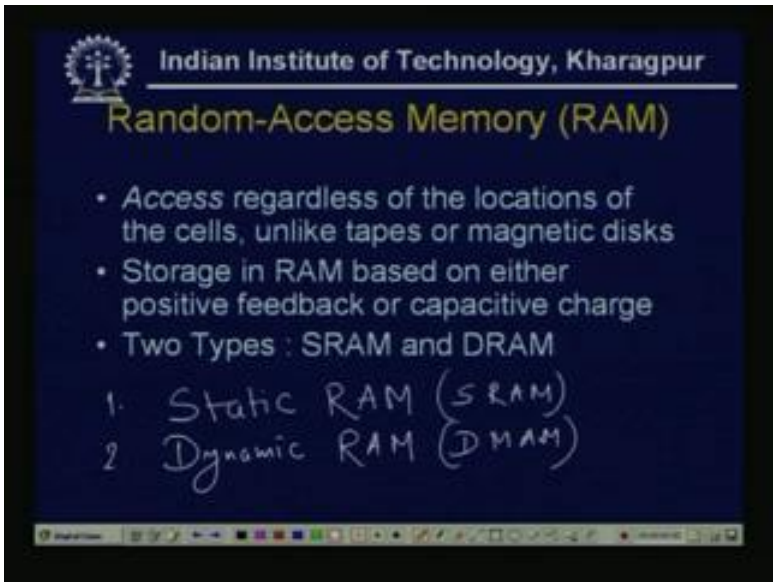
Digital System Design
Prof. D Roychoudhry
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur

Lecture - 31
Design of Memory Circuits

In last few lectures we have seen the design of arithmetic circuits, mainly there we have studied the design of adder, the different type of adders, the multipliers, the first multipliers etcetera. Now, today we will see another very important design structure, that is called the memory and that is one of the important circuits; that we need for any digital systems.

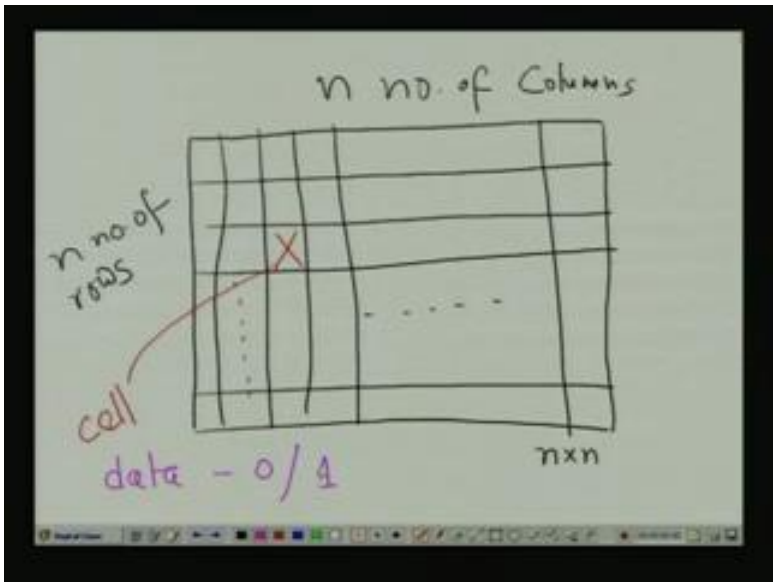
Today, we will read the design of memory circuits; now as already we have read the read only memory. Now, today we will see that the Random Axis Memory or the RAM, so first we define, what do we mean by the Random Axis Memories or RAM.

(Refer Slide Time: 01:47)



Normally, it access regardless of the locations of the cells, un likes tapes or magnetic disks, now storage in RAM based on either positive feedback or capacitive charge. But before this, the different types of memories, we see that, what is the normal structure of that memory, normally it is like this is a two dimensional array of cells.

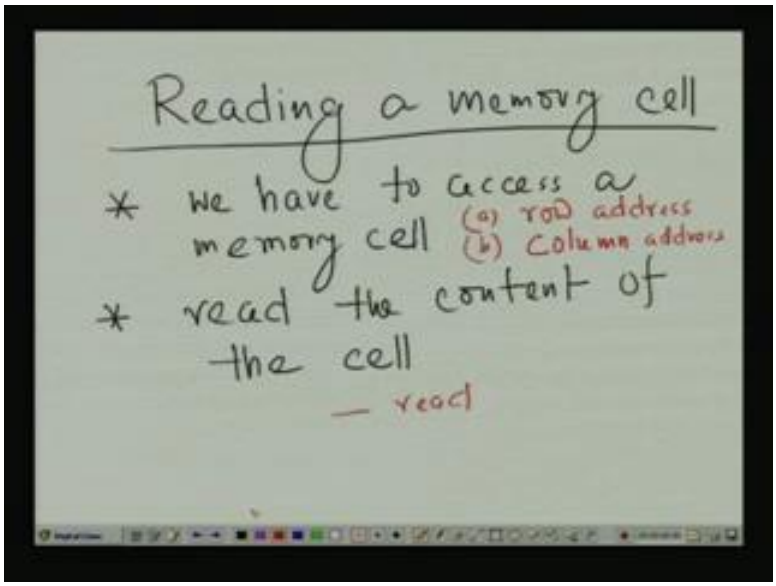
(Refer Slide Time: 02:32)



As if, we can treat that this is a two dimensional array of cells, so there are some rows, say n number of rows and similarly, there are n number of columns. So, here there are n into n number of cells, if we define that, so this is one cell, this is a memory cell. So, if I want to access, normally memory means, it stores a digital data; that means, either 0 or 1.

So, first thing is that they there are two operations in memory, one we call read operation; another is called the write operations. So, when we are considering the read operations; that means, some data are already stored in this cells. That means each cell contain either 0 or 1 and we want to access the content of that particular cell.

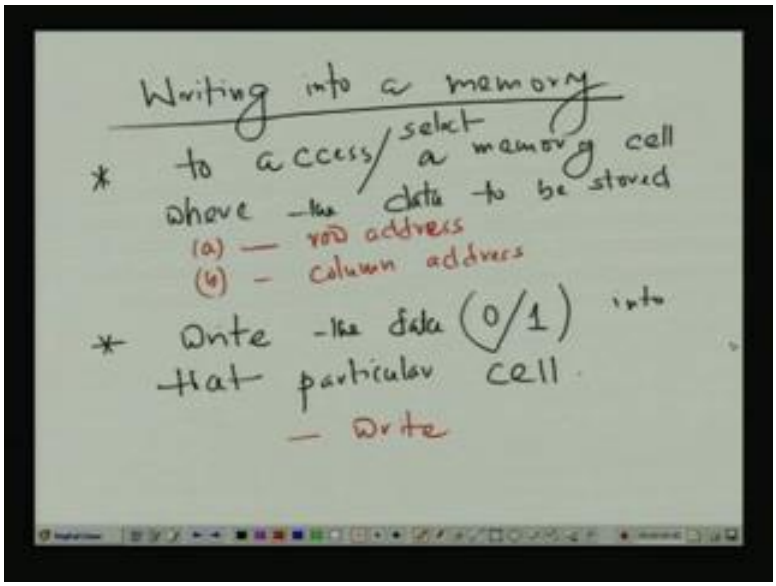
(Refer Slide Time: 04:42)



So, if I want to read; that means, for reading I have to read a memory cell. So, first thing is, we have to access a cell memory cell and then read particular content of the cell, so mainly these are the two steps to be followed. Now, to access a memory cell as just, now we have seen, that it is a two dimensional structure. So, to say this particular cell to access, we need to specify the row number and the column number.

That means in which, what is the cells position, so as it is a two dimensional array, so cell position means, normally that is the row and column and that is defined as a row address and the column address. So; that means, for each memory there is one row address and one column address. So, to access the memory cell, we need, first one is a row address or row index, we can call and one is a column address and here the operation is a read operation.

(Refer Slide Time: 07:00)

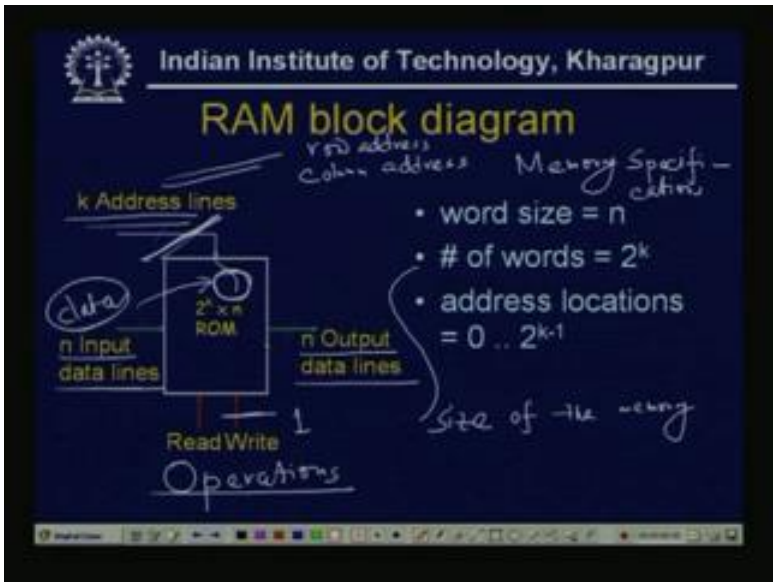


Now, similarly for writing; that means, when we want to writing a into a memory cell, memory cell again in which cell, where I want to write the particular data. So, for this purposes, again I need one to access and memory cell or better I can write here to select a memory cell, where the data to be stored. And then, the operation is write the data means 0 or 1value, 0 or 1 into that particular cell.

So, here also to access or select a memory, we have to specify a row address and a column address and here is a write operation means write the value 0 or 1into the cell. Now, when it is a reading, how we select the row address or column address and how we read the data, similarly when it is a write operation. Then, how we are selecting that, is the same row address and column address, but how we write, that is the mainly the memory operations.

So, mainly this is the storage in RAM based on either positive feedback or capacitive charge. Now, we can categorize the RAM, the random access memory into two types, one is called the static RAM, another is called the dynamic RAM. Normally, we specify these two as, one is the static RAM or the static random access memory or SRAM; another is dynamic random access memory or called DRAM. So, today's lecture will cover mainly the design of these two types of RAM.

(Refer Slide Time: 10:21)



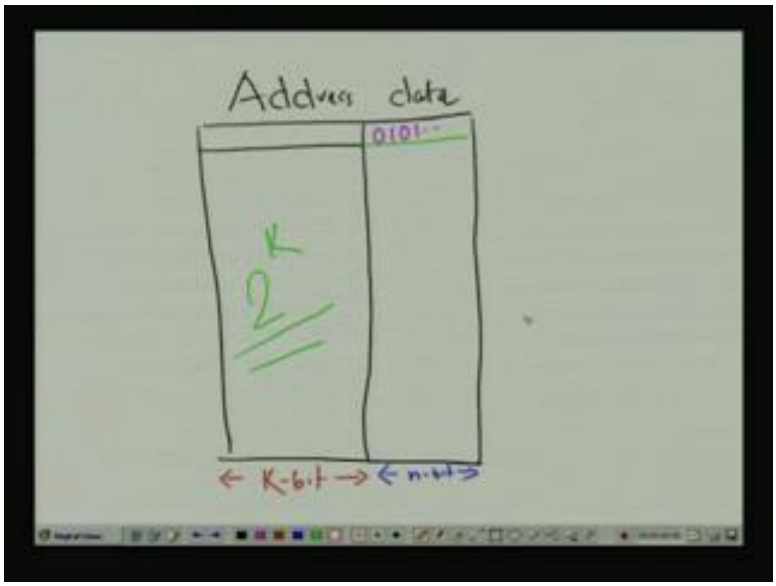
So, now we see the overall structure or design architecture of a RAM block diagram So, just now what we have seen, that the memory is treated as a two dimensional array of cells and each cell is actually it contains the data 0 or 1. So, it has some address lines the address lines, this is actually the row or column both, so these address gives, which particular cell, I have to access, I want to read or I want to write the cell.

So, mainly the row addresses, here the row addresses and the column addresses, so it specifies the cell to be read or the data to be written on the cell. Now, it has some n input data lines, means the data whether the data, if it is 0 or 1. That means, it is only 1bit data or it can be a n bit data, so then this is called a n input data lines.

Now, there are just now we have seen, there are two operations, either the memory, content is to be read or we want to write the memory. So, the two operations are read and write, so these are the memory operations and as it is a n input data lines. Normally, that is a n output data lines, because memory means, this is the storage, so n bit data, if it is stored, normally we want to get the n bit data, so that is why it is a n output data lines.

So, if we write the memory specifications, then it is a words size n means n bit data, number of words. That means, how many number of data are stored, there what is the capacity of the memory or sometimes, we call the size of the memory. So, these are 2 to the power k, so if it is a k address lines, then we have 2 to the power k number of storage.

(Refer Slide Time: 13:36)

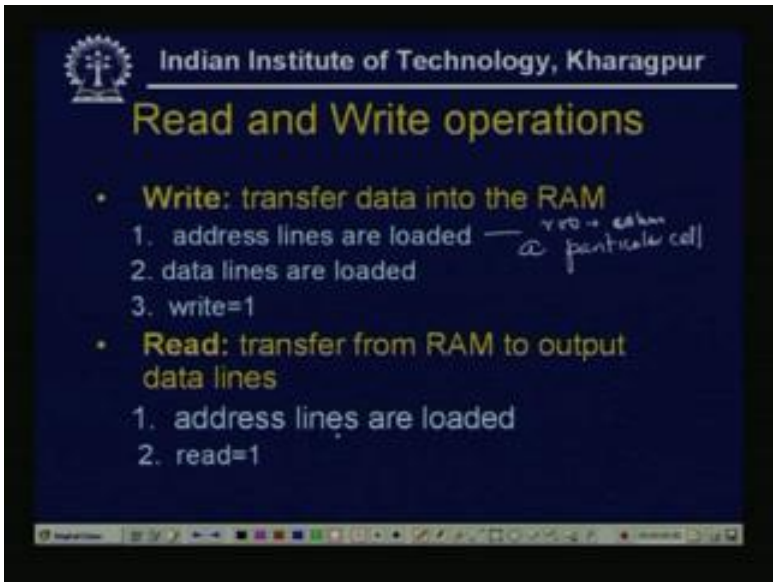


Again, this is similar to that of our table format, that what we have read in our read only memory, again what we can tell that as if, this is a table, where some data is stored, so it is the address part and it is the data part. So, if this is a k bit address, say this value is a k bit and then there are 2 to the power k such, addresses possible and that means, in this table, there are 2 to the power k number of data's are available. Say for each address, there is one data available, here the some 0 10 1and say this is n bit data.

So, 2 to the power k such n bit values are stored here, so this is my address and this is my data, just we can call this is a table or this is my memory, were the data's are stored. So, this is my word size n mean, that n bit data stored number of words 2 to the power k, just now what we mentioned this is the size of the memory. Sometimes, we call that this is my size of the memory, means 2 to the power k number of data can be stored.

So, address locations are normally 0 2 to the power k minus 1, means total 2 to the power k number of spaces are available, so just now what we have see, this is the k bit address utilize. So, 2 to the power k spaces and 2 to the power k, such n number of data are stored there.

(Refer Slide Time: 16:06)



Now, the two operations, already I have mentioned that normally memory, there are two operations read and write. So what write means, write is the transfer data into to the RAM, that means I want to store some value into the RAM. So, here the address lines are loaded, means I have to select the row and the column. That means in which a particular cell, then data lines are loaded, because I want to store, I want to write.

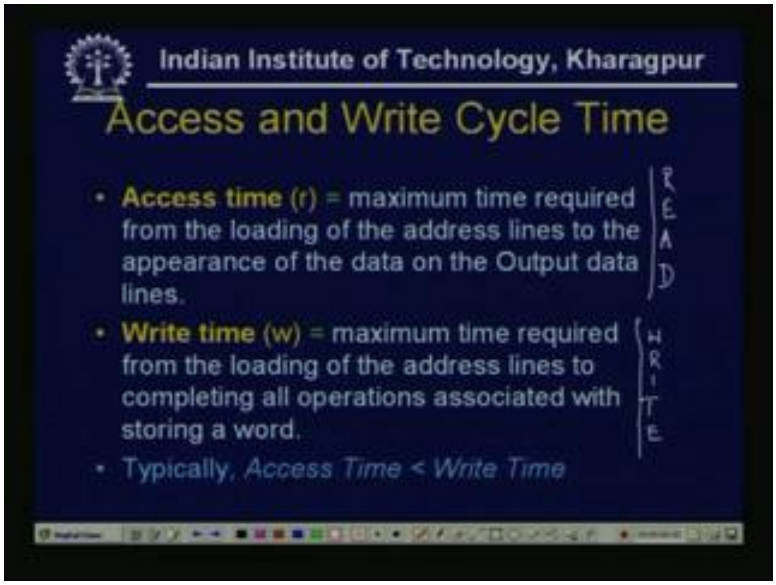
So, which value I want to write, if it is a n bit data line, then that particular n bit value is loaded into the data lines and then write equal to 1. So, if we see the design, here this write values, this should be this equal to 1, here the n input values, I have to load the data and before that, I have to select the k address lines, that means these three lines to be set.

This is a k address line it gives that in which particular cell, this particular cell have to be selected by this k address lines. Then, in detail value is loaded which is to be stored here and this is a write operation to specify that thing, this write line should be, so this is a write operation. Now read, so read is means transfer from RAM and because I want to read the content of a cell, so transfer from RAM to output data lines.

Now, similarly the address lines are loaded that means, which particular cell, I want to read, so similarly I have to specify the address lines. Then, read it is a read operation to specify that, now the read line should be made at 1, so read equal to 1and now we get the value to the data output line. That means that content of that particular cell or whose

address lines are given, the row index and the column index is given, then that content is at that output lines.

(Refer Slide Time: 19:20)



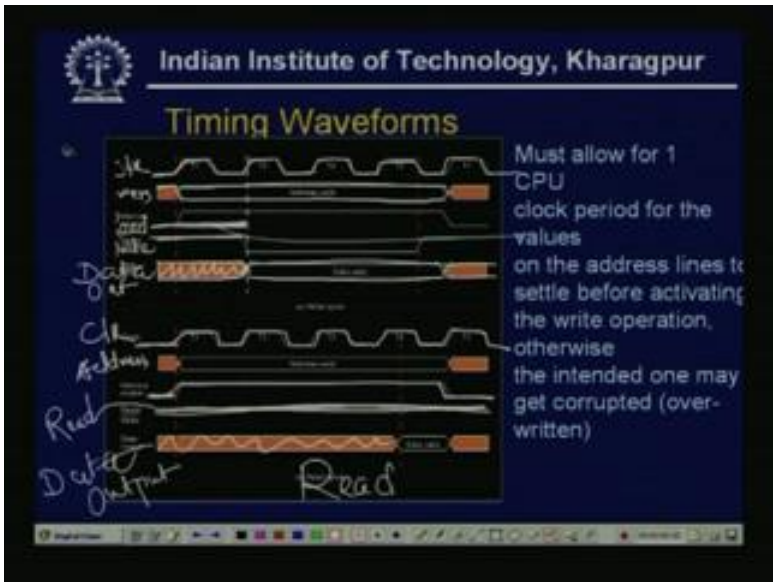
Then access and write cycle times, we define two other important things, regarding the memory operations, one is called the access time. Now, access time is the maximum time required from the loading of the address lines to the appearance of the data on the output data lines. That means, this is regarding the read operation of the memory, maximum time required from the loading of the address lines to the appearance of the data on the output data lines.

So, when we are reading a content of a particular cell, then first the part address lines of the particular cell are to be loaded on the address lines. The time difference between this loading the address lines and to get the content of the particular cell to the output data lines is called the access time r. So, this is the regarding the read operations, this is the regarding read, write time; obviously, this is regarding the write.

The maximum time required from the loading of the address lines to completing all operations associated with storing a word, again when we are writing as already I have mentioned, that again I have to select a cell. And selecting a cell means, I have to load a particular address lines, means I have to give a particular row address and the column address.

So, loading a address lines and then to complete all operations associated with storing a word, means I if it is a n bit value, then the full n bit value to be loaded, this whole time is called the write time. So, these are the two times, normally the access time is less than write time, that means reading time is lesser than the write time.

(Refer Slide Time: 21:44)



So, this is a timing waveforms, if we see that normally these allow must allow a for 1CPU clock period for the values on the address lines to settle before activating the write operation, otherwise the intended one may get corrupted or over written. So, thing is in brief, what we can tell that, we must give a 1CPU time or clock time to write the data or for the write operations.

Otherwise, in between that other data if it will not be stored, before that another data comes in. So, actually it will be a mess and that actual or we will get the wrong value. So, if this is my clock, say this is my clock and this is my address lines, this is my address data one address data is given. Now, this is the say this is memory read and write, so there are two lines the read lines and the write lines. So, this is read, this is write, clock this is the address.

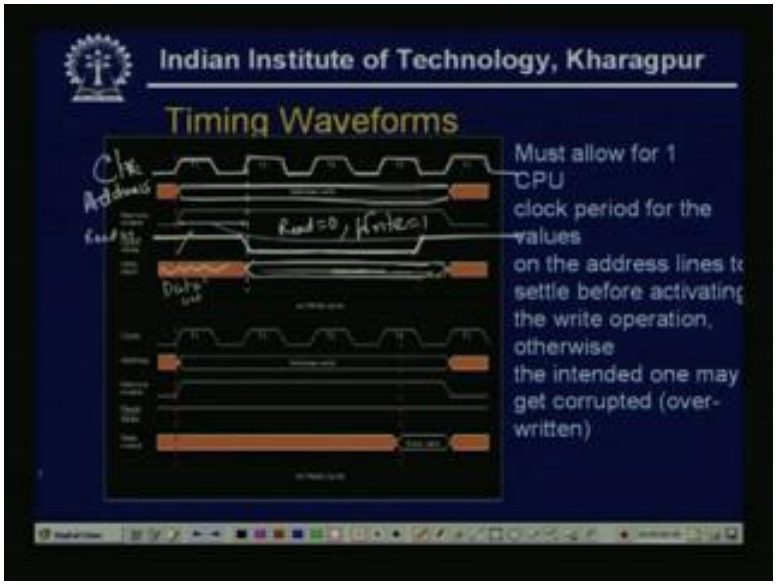
Now, if it is a write, it is read; that means, we are reading out, so this is my data out, so in this case it will be a data out up to this time it is a data valid, so this is read cycle. That means, when we are reading, what we have done, see this is my read is high means, read

equal to 1. This is read equal to 1and if the read is 1, then the data is going to the data output lines, so this is a read cycle.

So, again if we see that thing, if we see the clock, this is my system's clock, now again that address, so this is my clock, this is address, then this is memory enable, means this memory is selected. Now, if it is a read; that means, should be it is a read line, this is high, reading is high, read is 1, then this is my data output and this is totally read data, so this is my read cycle, this is my read cycle.

Now, similarly if it is a write, then instead of read, normally one read write one line, it should be low, read line should be low ,means write is high. So, if we see that, now if you see the timing waveforms for the write cycle.

(Refer Slide Time: 27:14)

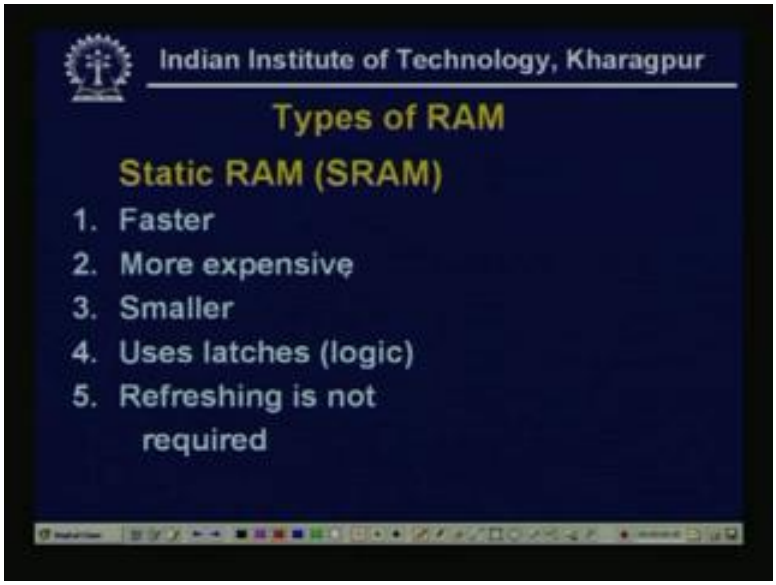


Then, again we consider that, this is my clock, then this is one particular address is selected, this is my address lines, now this read line becomes 0, means write is 1. Now, here this is memory enable, here it is read equal to high, read equal to 1, now, here in this phase read equal to 0, means it is a write equal to 1. So, in this clock cycle, it will be a write, so see here read equal to 1, so this must a data out.

Now, here as it is a write, now the cell selected here, the data will be stored, so here the data will be stored here, so this is a write stored. So, normally by the read write line, if read equal to normally, this is only one line, if it is high, it is a read, if it is a low. That

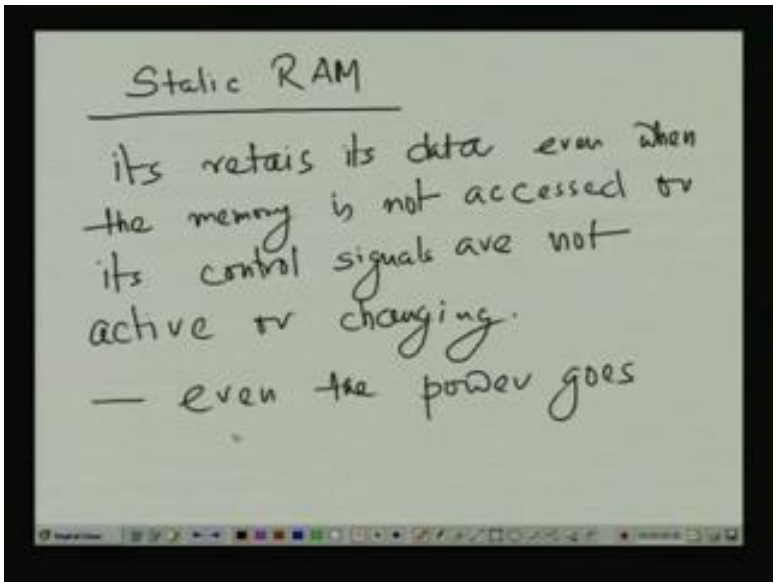
means, read equal to 0, means write equal to 1and then it will be a write and the both the cases read cycle and write cycle, the address lines will be selected.

(Refer Slide Time: 29:13)



Now, we see the different the two different types of RAM, what are their properties? So, first we defined what SRAM is and then this it is property will carry.

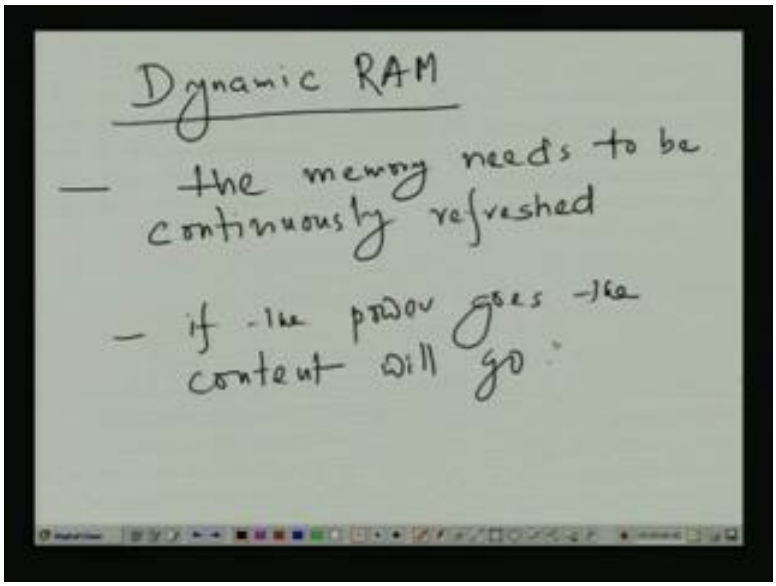
(Refer Slide Time: 29:31)



So, we can define the static RAM, it retains it is data even, when the memory is not accessed or it is control signals are active or changing are not active or changing. Normally, simply what how we define that, we tell that it is a static RAM, that even the

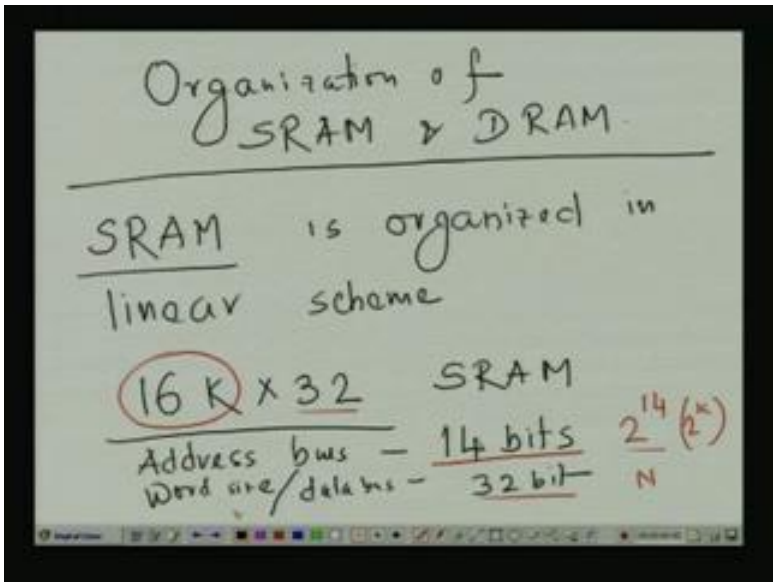
power goes, then also it will be there, that means the content will be retained in the cell. So, it retains its data content, even when the memory is not accessed or it is control signals are not active or changing or simply what we can tell that, even the power goes the content will be retained, so that is called the static RAM.

(Refer Slide Time: 31:33)



On the other hand the dynamic RAM, so simply what we can tell that, if the power goes the all the data will go or the data will vanish, then here the memory needs to be continuously refreshed and that is why it is called a dynamic. So, if the power goes, so content will vanish, the content will go, now there are some other differences also and from the organization point of view, there are some differences between the static RAM and the dynamic RAM.

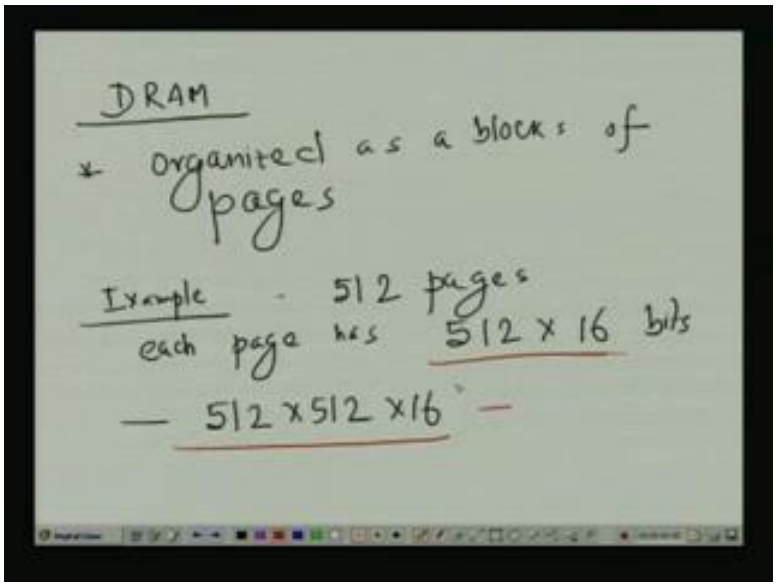
(Refer Slide Time: 32:59)



So, from the organization point of view organization of SRAM and DRAM, normally the static RAMS are always organized as linearly, it is called a that linear range of SRAM. It is or sometimes, it is called a SRAM is organized in linear scheme, what does it means, see if we have a 16 K by 32 memory, so 16 by 16 K by 32 SRAM. Then, the memory specification is it has 14 bits of address bus and what size or data bus, we can tell what size or data bus we can tell is 32 bits.

A 14 bits means 2 to the power 14is 16 K, so this is address, this is 16 K and here it is a 32, so this is my N, so actually this is 2 to the power K by N bit word size, so this is called a linear scheme. Always the static RAM is organized like that, so it is a simple structure; that means, the number of address lines directly gives, how many contents are there; that means, the address bus. Here that, it is a 16 K means 2 to the power 14; that means, 14 bit address lines and 32 bit data, that means the data bus 32 bit, so this is a SRAM organization.

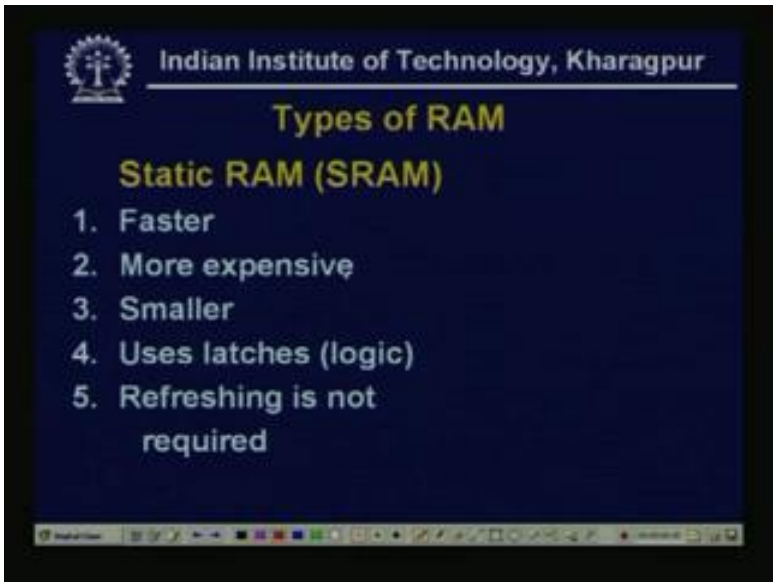
(Refer Slide Time: 36:13)



Now, if we see the organization of the dynamic Ram, so for DRAM, it is normally is a organized as page wise, this is organized as a blocks of pages So, it is a 512 by 512 into 16 memory, one example if we take that, if it is a memory is organized into 512 pages and each page has 512 into 14 bits. Then, the memories of I can write, this is a 520 by 512 into 16, this is the size of this is the organization.

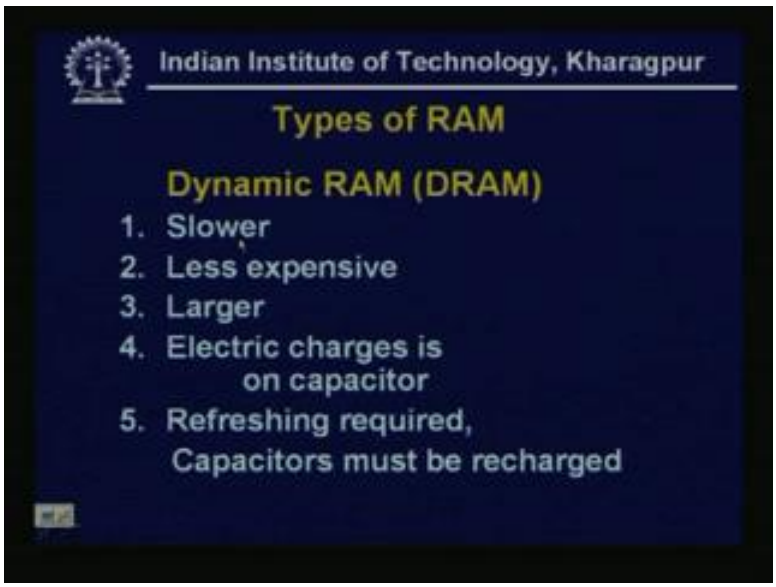
So, that is linearly organized mainly the address lines SRAM is organized as address lines and the data lines and here it is a blocks of pages. So, here if it is a 512 page and then each page contains 512 into 16 bits, then we can tell that this is the size of 512 into 512 into 16, this is the blocks of pages, in that way, normally the DRAM is organized.

(Refer Slide Time: 38:35)



So, now the properties the static RAM normally it is very it is faster, it is more expensive, because of it is power, even the power goes the content is retained, it is smaller and uses latches. Already, we have seen that actually the latch itself contains the data and refreshing is not required that is called the static RAM, that even the power goes the data content is retained.

(Refer Slide Time: 39:12)

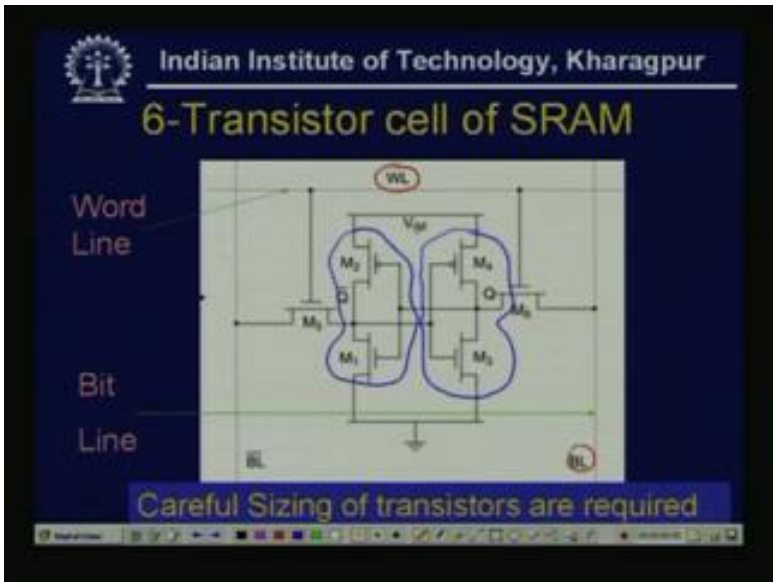


Now, dynamic RAM, so it is slower, less expensive, because if the power goes, then nothing is there, the data content will go, it is larger and electric charges is on capacitor

as it is a charges and discharges. If the power goes everything will go, so it is charges, Refreshing required and capacitors must be recharged. So, here the technique is that, when the power is there the capacitors are charged and that gives you the data 1 and if it is discharges, it gives the value 0, so mainly that is the principle of the dynamic RAM.

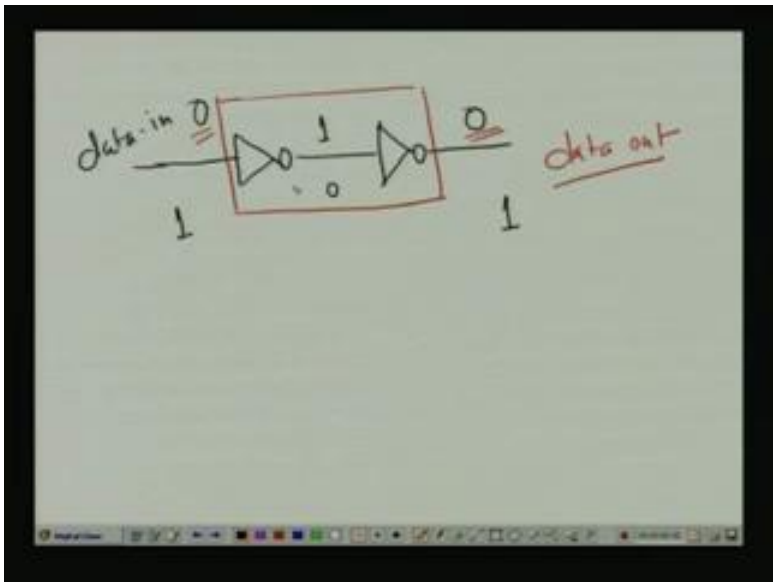
Now, one thing that both the SRAM and DRAM, there are they are of volatile memory; that means with time that the content discharges. So, in all are volatile memories, but this one is static; that means, if power is there, the content is there, if it is dynamic, then means if power goes, then the content goes. So, mainly these are the differences between the static RAM and the dynamic RAM.

(Refer Slide Time: 40:34)



Now, we see the actual architecture of a SRAM, first we see one SRAM, the static RAM, so there are different type of architectures, we consider here the 6 transistor cell of SRAM. So, this is a transistor SRAM, see here this is a word line, the word line is written as a W L, so this is the word line and this is the bit line, that B L. Now, there are 6 transistors are here, that M 1, M 2 this is a P type transistor, this is a N type transistor, so actually this is 1CMOS inverter, similarly M 3, M 4 this is another CMOS inverter.

(Refer Slide Time: 42:13)

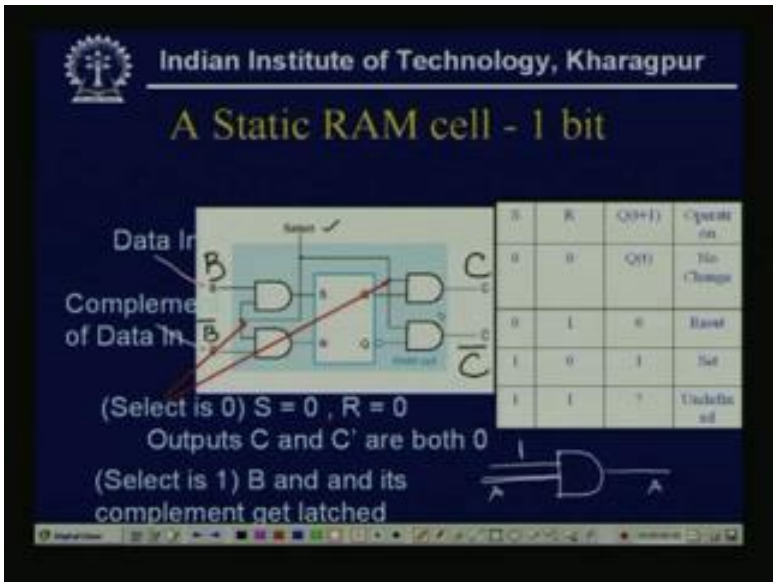


Now, if we remember the simple things see, what we mean by a storage or when we have consider, how we one memory element can be constructed. So, already what we have seen, that if say two inverters we store or connected here, see here the in data in is 0, so if it is inverter this will be 1, again if it is inverter this will be 0. So, now if we consider, this thing act as a black box, so actually my data in is 0, my data out is 0, this is my data out, so actually 0 data in 0 data out.

Similarly, if I give you one data in, then I will get one data out, because it will be inverter as 0, that means two inverters connected side by side, that is the concept of storage and see here this thing is there. So, actually this is nothing but a CMOS inverter M 1, M 2 and M 3, M 4 that is another CMOS inverter, they have connected side by side and Mth, M 5 and M 6 nothing but that gate is connected to the word line of the memory.

Similarly, that bit line is connected to sources of the two transistors M 5 and M 6 and it will work as that two inverters as if connected side by side, only the transistor design it is, so it is a CMOS design of 6 transistors cell of static RAM, it works like that. Already, we have seen that, this is mainly, these are for replenishes; the charge loss, the P transistors, the M 2 and M 4.

(Refer Slide Time: 44:42)



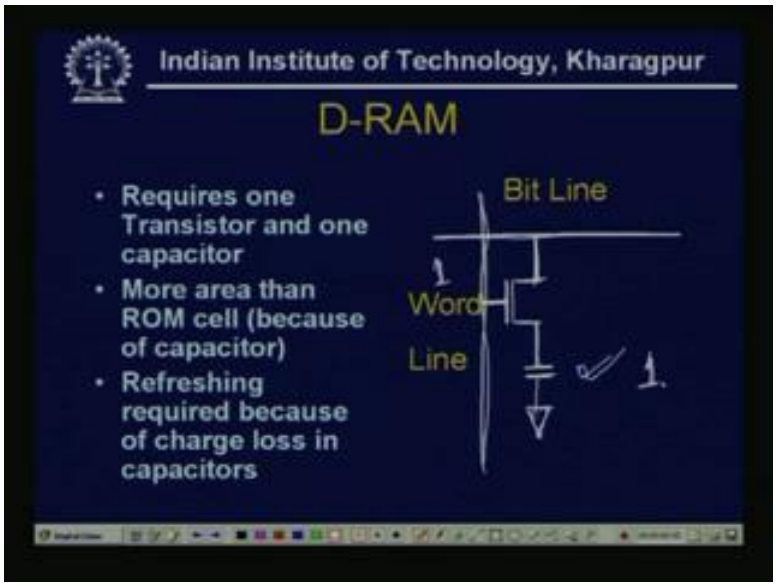
Now, is static RAM cell of 1bit; that is the CMOS transistor design, now if we see that simple SR flip flop this is nothing but a simple SR flip flop. So, here we have seen a data in and this is a compliment data in, so if we tell that this is B, this is B compliment. Now, this is the select line and if select is 0; that means S equal to 0, R equal to 0.

Already we have we know the functions o of SR flip flop, so output C and C dash are both 0, so if we see the truth table, that SR, if S and R both are 0. Then, operations that actually no change; that means, C and C dash that output both C and C compliment are 0. Now, if select is 1; that means, here B and it is complement get latched, so both B; that means, that up the upper AND gate, one input of the upper AND gate and another input of the lower AND gate and the other inputs are select that is 1.

So, it is actually this type of thing say 1AND gate 1AND gate, whose input is 1, so actually whatever I will give to the other input as it is a AND gate, so it will be A dot 1, it will be A. So, B will come to S, similarly B compliment will come to R, as these will be latched in the SR flip flop and that means, we will get, so if SR is 0 1, then it is a reset, it is 10, 1then, it is a set and if it is a 11,then it is a undefined state.

So; that means, see here it is a S is 0, this value is 0, it is 1. That means, the value is the value is latched. So, this is only 1, how 1bit of a static RAM can be constructed using SR flip flop.

(Refer Slide Time: 48:11)



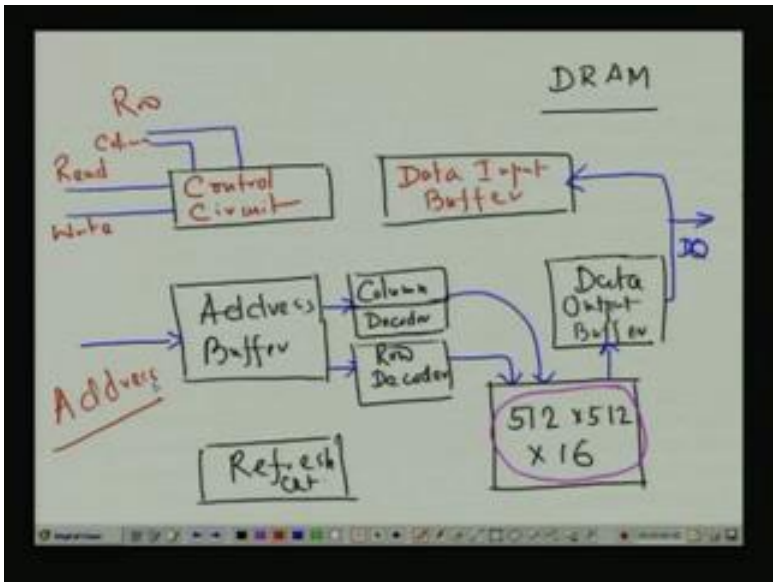
Now, it is a DRAM, so first we see that, only 1bit DRAM, a dynamic RAM, so requires DRAM requires one transistor and one capacitor, more than more area then RAM cell because of capacitor. As already we have mention that, static RAM is the concept is that, if the power is there, it is there because just now we have seen the architectural design, that we one using the CMOS transistor, another using the or SR flip flop.

So, there is the concept that if the power goes, it will be latched, so what data is stored it will be there. Now, here in dynamic RAM, if the power goes, it will goes means the concept is already I mentioned that, they want capacitor is charged and capacitor is charged means data content is 1. If it discharges, it is 0, some because of capacitors; more area is needed as we have seen that actually larger area needed for dynamic RAM.

So, refreshing require because of charge loss in capacitors, so the thing is very simple that say as if these horizontal row is the bit line and this is the is the word line. Then, one transistor actually this can be any type of connections, we have given one transistors here and as if the source is connected to the bit line, gate is connected to the word line. We know that, if the current flows means the actually this is ground.

So, if the word line is word line is high; that means, if it is 1, then whatever value is there. So, if it is bit line is also 1, then actually it will be charged, that means 1is stored, so what 1bit, that is stored here.

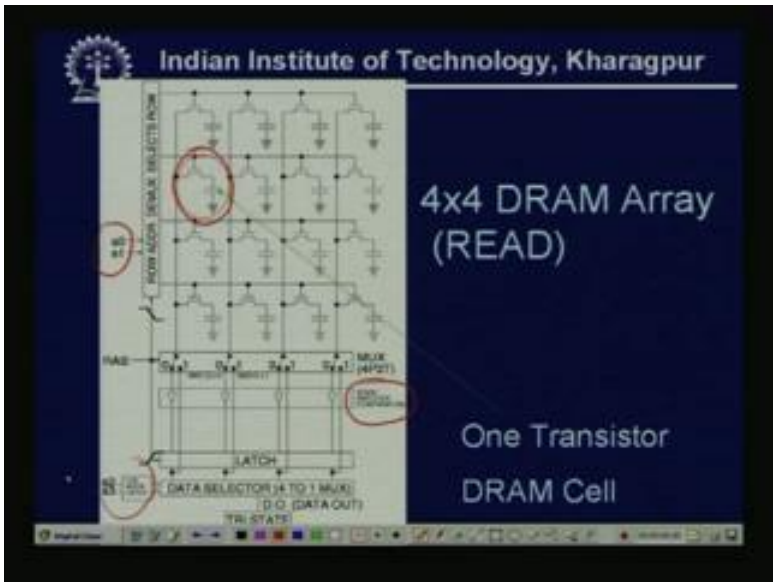
(Refer Slide Time: 50:49)



Now, before this we see the 1DRAM and block diagram level, one block diagram of DRAM, so it has some control circuit, these are the control circuit and it has four lines read, write the row address and the column address. So, the one is that read enable, write enable, this can be normally in can be one line and then row and the column addresses, so this is for row and this is for column.

Now, another will be the data input buffer, one will be the data output buffer, then this will be the DQ is the line for the data output. Then, there will be address buffers and that will go to column decoder, now if it is a DRAM, say of blocks of pages, so the example we have taken 512 by 512 into 16 bits. This is my memory, so from here the column address, column decoder, say it will go to the row decoder and the column decoder and some refresh circuits will also be there. Some refresh circuitry will also be there; obviously, there will be some address lines here, so this is the overall block diagram of DRAM.

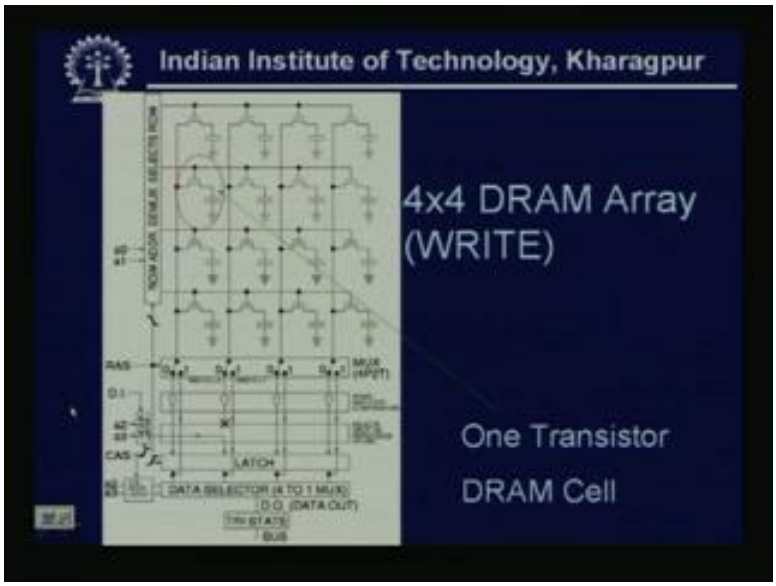
(Refer Slide Time: 56:39)



So, now we see that read operation of a 4 by 4 DRAM array, so it is a 4 by 4, just now we have seen for 1bit, it is nothing but a one word line, if we remember the structure, one word line, one bit line and only one transistor as the connector. So, just if we consider that one, if we consider that read, say there will be 4 by 4 DRAM, so 4 by 4 cells are there and an each connector, there will be one transistors.

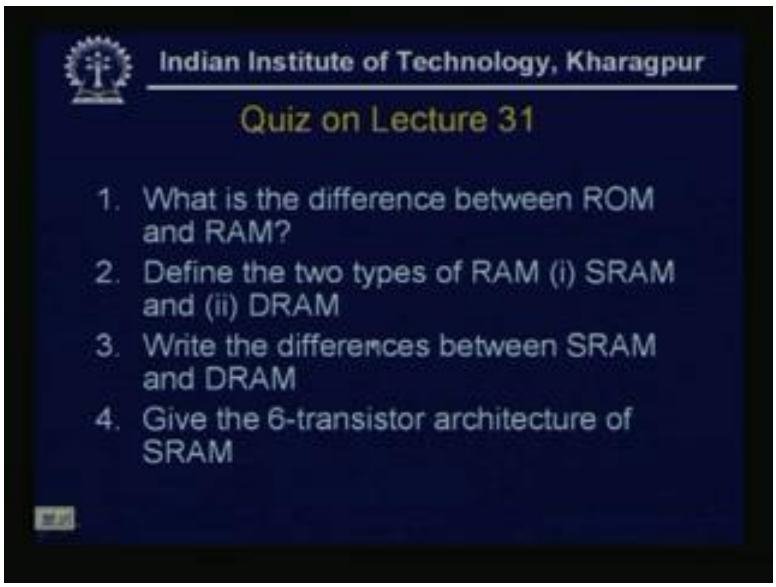
So, one transistor, 1DRAM cell, this is 1DRAM cell and these are the data selectors, that which one or see this is the address buses will come to here. Then, this is my sense amplifier comparator and accordingly it will be the row address and it selects row, it will select row and then we the columns will be selected and which cell to be read is there.

(Refer Slide Time: 58:31)



Similarly, this is the write cell exactly the same type of structure and so this is the example of a 4 by 4 DRAM array.

(Refer Slide Time: 58:38)



Now, this is the quiz on the lecture 30, one mainly the difference between the ROM and RAM. The difference between the static RAM and the dynamic RAM, already we discussed and the 6 transistor architecture of the SRAM. So, here we end the design of the memory circuits.

Thank you.