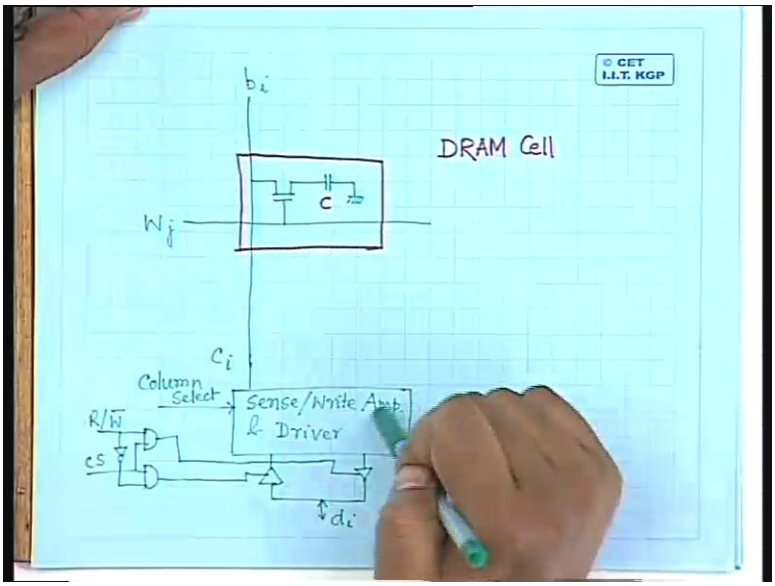
Digital Computer Organization
Prof. P.K. Biswas
Department of electronic and Electrical Communication Engineering
Indian Institute of Technology Kharagpur
Lecture No. # 21
DAM Architecture – I

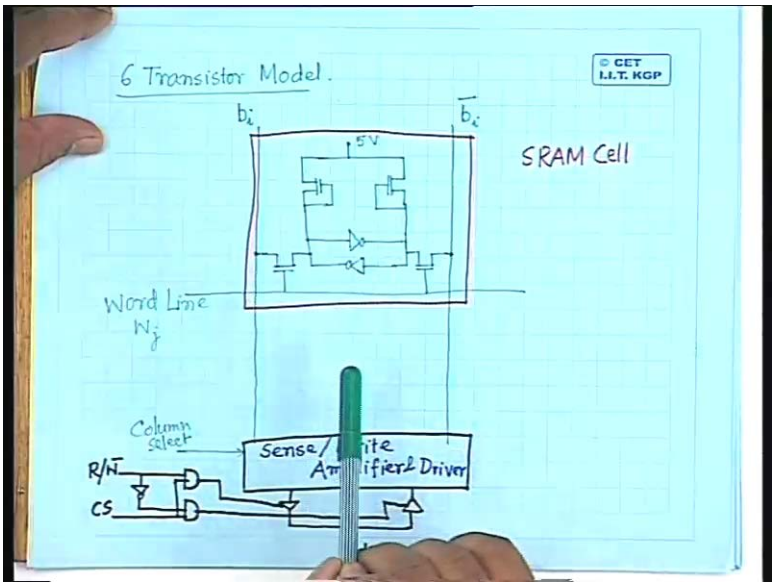
So, in the last class we were discussing about the dynamic RAM or DRAM cells and we have said that a DRAM cell is a simple circuit like this where you have just a charging capacitor and a charging discharging part. So depending upon what voltage is stored on the capacitor, you would store either bit 1 or a bit 0. So for storage or for retrieving the value that is stored on the capacitor you have to enable the corresponding row, so the entire row of cells will be selected.

(Refer Slide Time: 00:01:29 min)



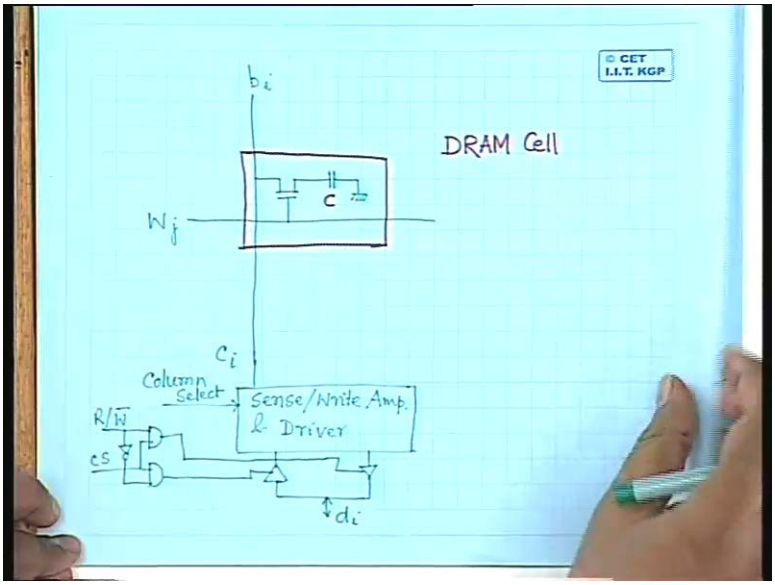
Then from that row you have to select a particular column, so one or more columns using the column select in which the control comes here and unlike in case of static RAM where we have said that in case of static RAM, the basic cells consists of a latch, I have that figure. So this is the static RAM cell and since the basic cell consists of a latch, so once we store value the value will remain until and unless it is over written by a new value.

(Refer Slide Time: 00:01:54 min)



Whereas in case of a dynamic RAM, since the bit 1 or bit 0 depends upon whether the capacitor is charged or discharged over a long period of time or even when you read the content of a particular capacitor, if the capacitor was previously charged then because of reading operation the capacitor will be discharged.

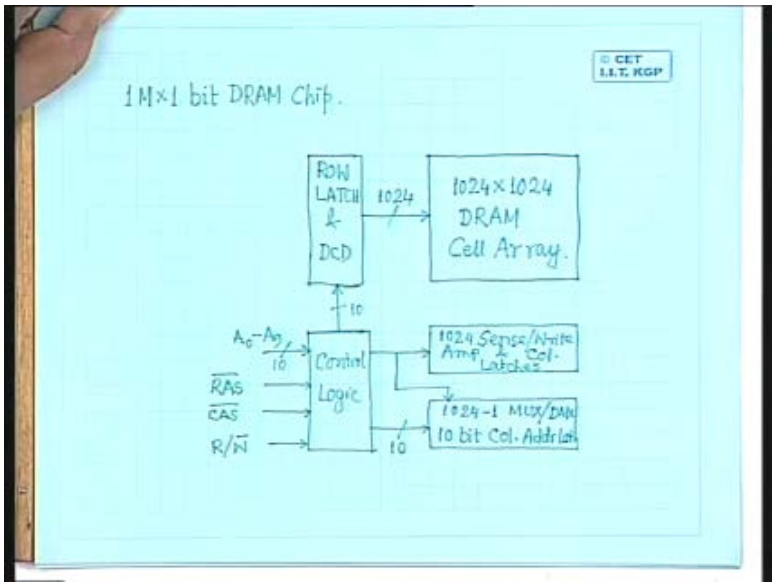
(Refer Slide Time: 00:02:05 min)



And not only that after storing certain value even if you don't read the dynamic RAM for a long period of time then the capacitor will get discharged because of leakage current. So in case of dynamic RAM, we must have a provision of refreshing the content of capacitors at regular intervals of time, so which is taken care of by this unit that sense right amplifier and driver.

So even if you read it, after reading operation you have to recharge the capacitor after amplifying the value that has been sensed on the capacitor. Even if you don't perform any read operation even then it will be the responsibility of this circuit or a control circuit, in addition to this I will come to dynamic RAM organization. There are control circuits, responsibility of the control circuit will be to write, to read and write the content of different rows in the dynamic RAM cell array at regular intervals of time and that has to be taken care of by an additional control unit which is given along with the DRAM. So coming to the DRAM organization, the organization of a DRAM is something like this. Let us consider that we have a DRAM chip having one M location and every location consisting of one bit.

(Refer Slide Time: 00:03:46 min)



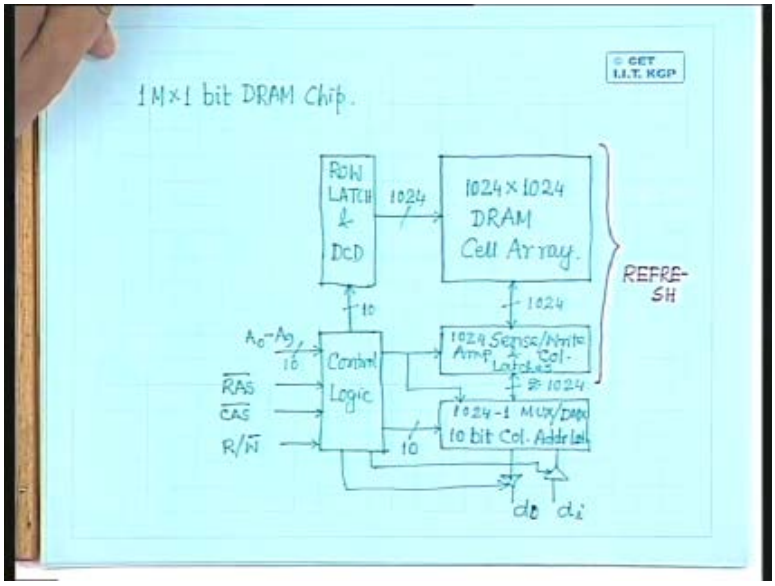
So 1 M by 1 bit DRAM chip. So because it is 1 M by 1 bit so the typical organization of the RAM cells can be a two dimensional organization. So the kind of organization of RAM cells that we will consider is something like this. Let us assume that we have a two dimensional organization of RAM cells which will have 1 k, rows every row consisting of 1 k number of cells. So we will have 1024 by 1024 DRAM cell array. Now others so, because it is 1 M locations and every location consists of 1 bit that means the number of address lines that we have to have is 20 address lines. Now another consideration which affects the design of DRAM chips is that the market of dynamic RAM cells is very competitive and once you design the DRAM chip then once the RND cost is made up, later on the cost of the chip depends upon what is the fabrication cost and the silicon cost. And silicon cost is very less because we hardly use any silicon image chip. So what decides the cost of the, it is not only for DRAM, the cost of any chip is what is the fabrication cost. And fabrication cost almost increases linearly with the number of pins that you have.

So one of the recent constructions is to reduce the number of pins which should be true for all other chips but it is specifically important in case of DRAM because DRAM has very high competition in the market. So, all the manufactures try to reduce the number of pins that you have in the DRAM. So in this particular case because the number of pins needed for just feeding

the address lines is 20, so what manufacturers prefer is that instead of having 20 address lines let us have 10 address lines. And we will feed row and column addresses because we have seen in case of other chips also, even in case of SRAM chip that the address bits that you provide that is internally divided into two components, one is row address other one is column address. Row address go to row decoder, column address goes to the column selectors. The row decoders activate a particular row, out of that row the column selectors decide or take out the number of columns which are needed.

Similarly in this case, but in that case we don't have any row column multiplexing but in case of dynamic RAM because of competition, manufacturers have decided to go for row column multiplexing. So we will assume that there are 10 address lines over which 20 address bits will be provided and the bits will be time multiplexed. First you give the row address which is 10 bits then you give the column address which is again 10 bits. So we will have two additional components, one of the components we call as row latch and decoder because I have to have row address decoder. Then among the other additional components we will have because once you give a row address, the entire row of the cells is selected either for read operation or for write operation and in this case every row consists of 1 k number of cells. So here I need 1024 number of sense right amplifier and column latches. That means the entire 1024 number of cells belonging to a particular row will be latched into this 1024 number of column latches. From this 1024 number of columns, I have to select out one particular bit which is to be done by the column selector or a multiplexer.

(Refer Slide Time: 12:10)



So I have to have another component which will be 1024 - 1 multiplexer or demultiplexer from 1 - 1024 that will be needed for the writing operation. In addition to this I have to have a 10 bit column address latch and here I have to have controller and these are all additional circuits which are not needed in case of static RAM array. This controller will accept 10 bits of address A_0 to A_9, in one case these 10 bits will contain row address, during other time instant these 10 bits will contain the column address. So initially the 10 bits of row address has to go to the row address

latch, here also we have 10 bits and at later instant of time the 10 bits has to go to column address latch, when you feed the 10 bit column address. In addition to this there has to be control signal given to the sense write amplifier and column latches. We will also have control signals coming to the 10 bit column address latch. Now inputs to this controller, so this is the control logic.

In addition to this 10 bit address pins, we have to have some additional signals to identify that when you feed the row address or when you feed the column address. So I have two additional signals, one is called RAS or row address select. Normally this is given in the imparted form so we have RAS complement, the other control signal is CAS or column address select bar. In addition to this we also have to have the control signal of the read or write bar. So when this 10 bit row address comes to this row latch, row address latch and row address decoder this will give 2 to the power 10 that is 1024 number of row select lines. So depending upon and one of these will be activated at a time and depending upon which line of this 1024 lines is active, the corresponding row in this 1024 by 1024 DRAM cell array will be selected either for read operation or for write operation. That means this entire row will be available to this 1024 sense right amplifier and column latches.

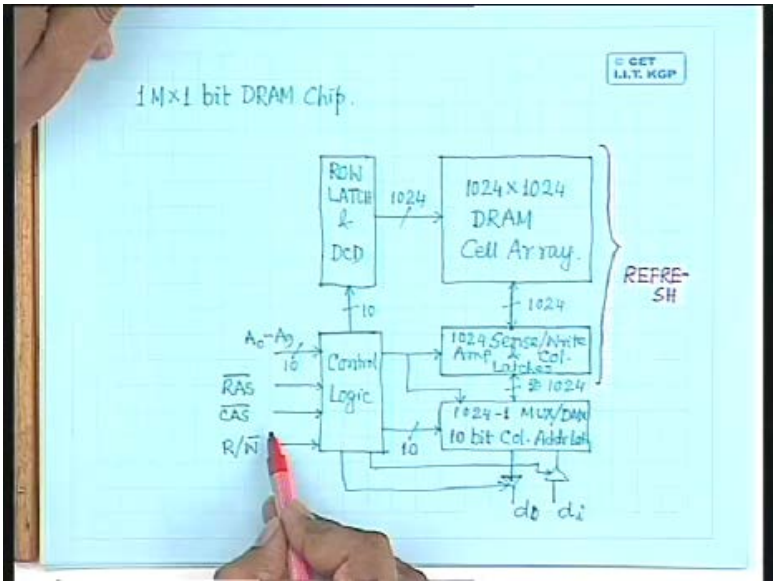
So here also we have 1024 number of lines, from this the same 1024 lines is also available to this multiplexer demultiplexer as well as 10 bit column address latch unit and from here I can have inputs and outputs. So if it is output that has to go through a output buffer, if it is for input purpose then demultiplexer will be active which gets an input from the input buffer. These buffers are to be activated from the control logic only depending upon whether the signal is a read signal or a write bar signal. So this is say d_i ith output line and this is the output line and this is the input data. So this is the entire unit of a RAM chip and the RAM chip organization is something like this. So whenever you give the row and row address select, these two control signals together, the 10 bit data which is available on this row address lines that go to row address latch and decoder.

Depending upon the decoder output, one of the rows consisting of 1024 number of cells will be active which comes to this sense right amplifier and column latches from where it is available to multiplexer demultiplexer unit. now when you give the column address, simultaneously the column address select this has to be active following which the column address 10 bits is latched into this 10 bit column address latch and depending upon what bit you have addressed, what 10 bit combination is latched in this column address latch unit, the multiplexer and demultiplexer unit will act accordingly and depending upon that you have one bit at the output or one bit is selected for writing into a particular part. Now you find that as I said, I also have to have a provision of refresh. So the refresh operation in this case can be done like this, if I give only the row address select input with the row address, I don't provide any column address. In that case the operation that will be performed is read the particular row, amplify the content and write it back. So then the unit which is active is only this part in the refreshing unit, this part does not come into picture.

So only this is the portion along with the row latch and decoder and the control logic that takes part in the refresh operation. This multiplexer demultiplexer unit are 10 bit common address latch, this unit does not take part in the refresh operation and you find that this unit is active only

when the CAS bar is active that means we provide a column address. So data from one of the cells will be available on the output data line, when you give the proper column address along with the column address select or data from the output will be written into one of the cells in this DRAM cell array again when you give the column address along with this column address selected input. And this is the line which is connecting to the external volt.

(Refer Slide Time: 16:48)



So as we have seen that in case of static RAM, only when you give the chip select then only the data lines gets physically connected. If your chip select is not active then the chip does not give any output on the data line neither it accepts any input from the data line. The same operation is done by column address select in this case. So in case of dynamic RAM, we don't need any additional chip select input. It is a column address select control input that acts the purpose of chip select also. So in this case your operation will be, you have to provide RAS bar followed by CAS bar when either read or write operation will be performed along with refreshing. If you provide only RAS bar, no CAS bar then read and write operations will not be performed on this DRAM cell array but the operations which will be performed on the DRAM cell array is simply refresh operation.

Say as I said that if I don't perform any… One case is in case of DRAM cell if I perform any read operation, the capacitors are going to be discharged. So I have to recharge it which is done by the sense right amplifier portion but even if I don't perform any read operation, even then the capacitor gets discharged if it is left untouched for a long of period of time. So I have to perform some refreshing operation and because there is no read or write operation to be performed on this that means CPU does not want any data from the DRAM, so CAS bar will not be active. So I have to have an additional controller in the DRAM unit. What this diagram I have given, it is only for a DRAM chip and number of such DRAM chips are to be connected together to give you a DRAM board. I will come to that organization later.

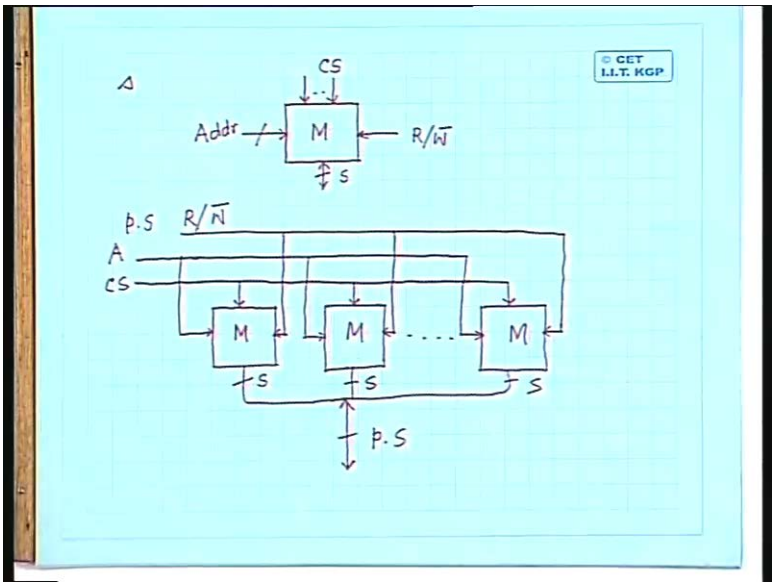
In that DRAM board, I have to have an additional control unit which will generate only this RAS signals along with the row addresses coming sequentially. So it will be responsibility of that controller that even if the DRAM is not read or written into, the additional control unit will provide the row addresses along with the RAS bar signal at regular interval so that different DRAM locations are refreshed. But it does not generate any CAS signal because CAS signal is only needed, this column address select only needed when the CPU wants to read some data from the DRAM or the CPU wants to write some data into the DRAM because CAS signal is connecting this to the external bus. So this is the logical organization of a DRAM chip.

I will come to the DRAM board organization or the SRAM board organization now. So far what we have discussed is individual DRAM chips or individual SRAM chips. Now when we design a memory module or a memory board depending upon the system requirement, I will have to connect more than one chips in proper fashion so that I can have sufficient number of locations, memory locations not only that I can have sufficient number of bits per location. Say for example in this case every DRAM chip will give you one bit per location but if I want to organize this, so that every location should contain 16 bits, I have to connect so many DRAM chips in parallel. Not only that this is also true in case of SRAM. If I have an SRAM chip which is a nibble organized that means every location in the chip contains 4 bits but maybe my bass width is 16 bit.

So for proper or efficient data transfer, I need that every memory location should contain 16 bits so that all the 16 bits can be read or written by a single memory access operation. So in such cases I have to combine or I have to connect 4 such chips in parallel so that for every read or write operation I get 16 bits together, I can read out 16 bits from every location or I can write 16 bits in every location of the memory. and I feel you have already done that in your previous course that how to connect different memory chips for memory expansion. Say for example if I have a memory chip where every location contains let us say s number of bits. By now it should be clear that if I have an static chip, static RAM chip then the logical diagram of a static RAM chip is something like this. It has to have a number of address lines.

The address lines are internally divided into row address and column address part. So I have to have a number of address lines then I have to have a read write bar signal, it will have one or more chip select lines. In some chips I can have a single chip select line, in some other chips I can have multiple number of chips select lines. and I have to have a number of input output data lines, so that is S. So using such a type of memory chip, if I want to have say p into s number of bits per location then what I have to do is I have to connect p number of such chips in parallel. The chip select inputs of all the chips will come from a common chip select input and this is the one which decides whether your chips are connected in parallel or the chips are connected in serial fashion. So this is a common chip select input which selects all these chips simultaneously.

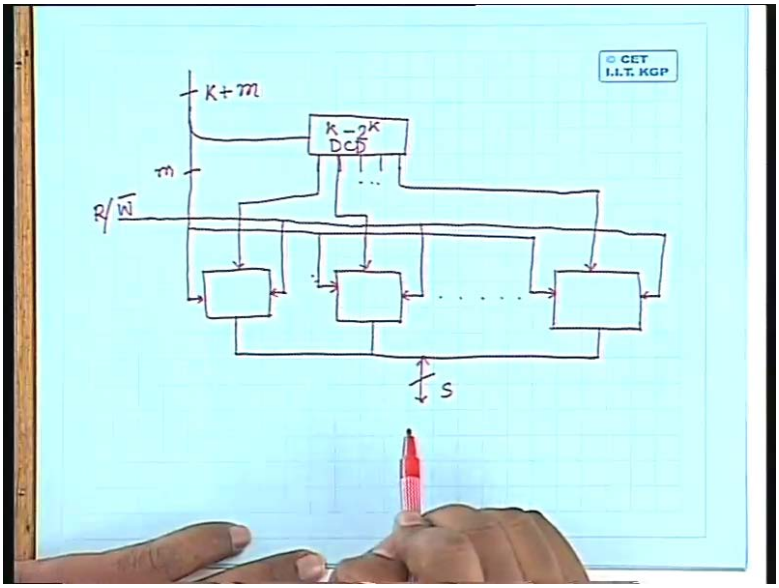
(Refer Slide Time: 00:21:47 min)



Then again because we are connecting them in parallel, so zeroth address in this chip must be same as zeroth address in this chip, it should also be same as zeroth address in this chip. That means every location in every chip should conform each other. So what we have to do is we have to have common address lines which should go to all the chips in parallel. So these are the address lines. I should also have common read write lines, so this is read write bar. In case of data lines what you have to do is they have to be concatenated. For each of the chips we have S bit data lines, when they are concatenated they give p into s number of data lines. So this gives you the simple parallel connection of p number of chips to expand the data bus by factor of p. So this is a simple parallel connection.

Now if we want that we don't want to increase the number of bits for the data bus but we want to increase the number of locations in the memory then instead of going for such a kind of connection, what you have to do is we have to play with the chip select lines. So assuming that every chip has got say m number of address lines and the number of locations that we want to have is 2 to the power k into m. So what we have to do is from the CPU, we have to give an address which is k plus m number of bits, m number of bits per chip gives you 2 to the power m number of locations. I want to connect 2 to the power k number of such chips in sequence. So total number of address lines becomes k plus m, out of this k plus m address lines, the higher address k bit addresses, I give to a k to 2 to the power k decoder. So I will have 2 to the power k number of output lines from this decoder. I connect 2 to the power k number of chips in such a way that chip select of every chip will come from one of the decoder outputs.

(Refer Slide Time: 00:26:34 min)



So maybe chip select of this chip is coming from this decoder output, chip select of this chip will come from this decoder output. So similarly chip select of this chip will come from this decoder output. Now once a particular chip is selected, a particular location in the chip can be selected independently. So in that case what we will do is we will provide the address lines to all the chips in parallel, these m bit address lines. So whenever this chip is selected, this m bit address lines, the m bit address is effective only for this chip. Though, the address is going to all other chips but for other chips that is not effective.

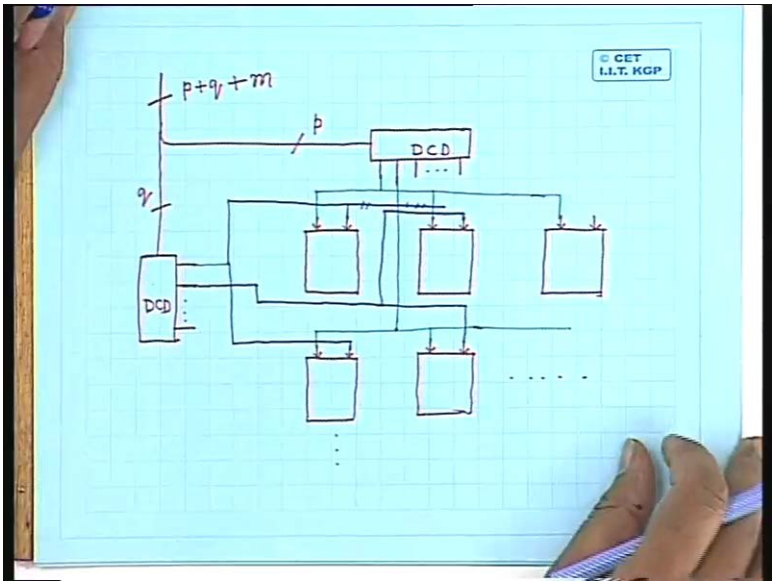
Similarly when this is one selected, this m bit address is effective only for this chip. It is not effective for other chips because others are not selected. What you do for read write lines? Read write lines again can be common because that depends upon the chip select output, so I will simply connect read write lines also in parallel. This is the read write bar signal and for the data lines what you do is you simply short all the data lines. Now in this case, it will not be concatenation because concatenation is needed when I wanted to increase the number of bits per location. In this case the number of bits per location will remain the same, only the number of locations has to increase. So you simply connect all the corresponding data lines together and the number of bits on this data line, you get s number of bits. And this connection is possible because we said, when you discussed about the memory cell that for every input and output, for every cell we have buffers and it is because of the buffers we can connect them together. If there was no buffer, such a connection would not have been possible. In that such case what we would have to go for is again a multiplexer demultiplexer there. So this kind of arrangement whether this one or this one, they are linear arrangement.

Now if I want to have a situation that I want to increase the number of bits per location, at the same time I also increase the number of locations in the memory board, what I have to do is I have to use combination of this and this. So once I have this parallel combination, you can think that this entire assembly looks like a memory module having, if m is the number of address lines 2 to the power m number of locations, every location consisting of p into s number of bits. This

whole thing I can consider a memory module of 2 to the power of m locations, every location consisting of p into s number of bits. So what I can do is I can simply replace each of these units by this module. So in such case I can both increase the number of bits per location, also at the same time I can increase the number of locations in the memory module. So here you find that in this organization, I have made use of only one chip select lines and the decoder that I have used is k to 2 to the power k decoder.

Now if this k is large in that case the decoder complexity will also be very large but we have already seen that when we discussed about the row column decoder, in case of a memory chip that only purpose of dividing a single decoder into two decoders is to reduce the decoder complexity. Similar concept can also be used in case of memory board. So if the value of k is quite large in that case to reduce this decoder complexity, what we can do is we can break this value of k, the number of bits again into two components. For each component I will use one such decoder, so both the decoders will give you some outputs. These outputs will now be used to activate the chips. So I can have a situation like this that instead of having say k plus m number of address lines, suppose this k is broken into p plus q number of address lines plus m number of address lines per chip.

(Refer Slide Time: 00:32:11 min)



Now I will have two decoders, one decoder will be say p to 2 to the power p decoder, here I decode p number of address lines. Out of this q address lines can go to q to 2 to the power q decoder, so this is one decoder, this is another decoder. Here I will have total 2 to the power q number of outputs, here I will have total 2 to the power p number of outputs. The chips I will organize in the form of a two dimensional array. So like this, the chips will be organized and here I assume that every chip will have two select input lines, one is chip select one, other one is chip select two.

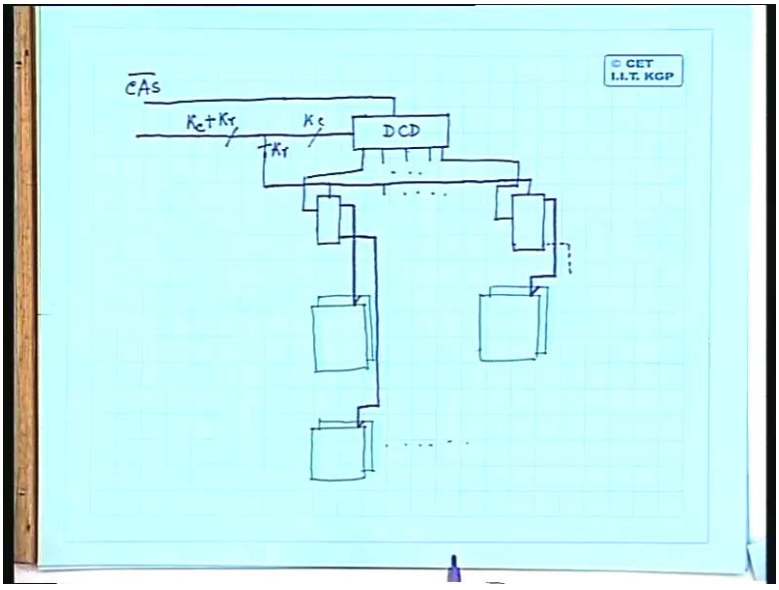
So now the chip select inputs can be used like this that every decoder output will activate one of the rows. So I connect this here, this here, this one here (Refer Slide Time: 34:35). Similarly this

decoder output can activate the next row and this way it can continue. For the other select input, we will use the output from the second decoder. So I can put it this way that this decoder output will activate this chip, this same decoder output will activate this chip. This decoder output, the second decoder output will activate all the chips in a particular column, so it will continue like this. So now we find that only when this decoder output is active and this decoder output is active then only this chip is active, other chips are not selected.

Similarly when this is active and say this one is active, the same decoder output is active then this chip will be selected. So this way now I can divide the memory cell array in two dimensional arrays like this and the advantage is the complexity of the decoder is reduced to a great extent. Regarding your address lines and other address lines, other m bit address lines they will go parallely to all these chips, read write signals will go parallely to all the chips and the data lines will be connected together. So unlike in the previous case where the memory chip array was one dimensional array, in this case the memory chip array is a two dimensional array and accordingly I have two decoders, one of them may be called as horizontal decoder, the other one may be called as vertical decoder.

Now this same concept can also be extended in case of dynamic RAM chips. So if we want to have a dynamic chip, RAM chip array, we have said that in case of dynamic RAM chips, the function which is done by chip select in case of static RAM chips, the same function is performed by CAS bar in case of dynamic RAM chip. And it is the dynamic RAM chips that can be used in a multidimensional array and the different array components can again be selected through a number of decoders. So in case of dynamic RAM chip, the organization can be something like this. Firstly we have to have a CAS bar signal, if whenever the RAM chip is connected to the external data bus.

(Refer Slide Time: 00:37:53 min)



Now this CAS bar signal in case of dynamic RAM can be used to activate a decoder, this is a decoder. The decoder will have a number of address bits as input and it will obviously generate

some outputs. So we find that this decoder will be active only when you get the CAS bar signal. Again if we want to go for a multidimensional array then what can be done is these outputs of the decoder can be used to activate other decoders and these decoders in turn will get a number of other address lines. So if I assume that here I have a number of address lines say K_c plus K_r where K_r is the number of rows, 2 to the power K_r will be number of rows and 2 to the power K_c will be number of columns. So, total number of chips that we can address is 2 to the power K_c plus K_r. So out of these address lines, say here we give say K_c address lines and K_r number of address lines are connected to each of these decoders, so like this it will continue. Then we will have a number of memory chip arrays.

What I can do is output of this decoder can be connected to CAS bar of this array. Similarly output of this decoder, other output can be connected to CAS bar of this memory chip. Similarly this can be connected like this, this goes to CAS bar of this chip. Similarly this output can go to CAS bar of other chips. So you find that when this memory chip will be selected, this decoder output is active and also this decoder output is active. So if I say that this is zeroth output of this decoder and this is also the zeroth output of this decoder then when both K_c and K_r, both of them are zero then only I am selecting this particular chip. Now I have given K_r number of bits address bits to this particular array this particular decoder, so this output may be 2 to the power K_r th output, 2 to power K_r minus 1 output. So this select, whenever this is active that indicates that K_c is zero and all the bits, all these K_r bits are equal to 1 then only I am selecting this particular memory chip. And if I assume that every location in the dynamic RAM array, in the dynamic RAM chip will be one bit and I want to have say s number of bits then what I have to do is I have to connect s number of such dynamic RAM chips in parallel. So it will continue this way and CAS bar of all these dynamic RAM chips will be connected in parallel.

So whenever I do this and otherwise your read write bar signal or the remaining m number of address signals, similarly data bus all of them will be connected in the same way as we have done before. So the selection mechanism which is only important is the way you select a particular dynamic RAM chip through the CAS bar signals. So along with this, as we said that we have to have something to generate and similarly RAS bar signal which can also go parallel to all these chips because until and unless CAS bar of a particular assembly is active that particular unit will not be selected. So RAS bar signal can go parallely to all the chips, read write bar signal can go parallely to all the chips. Then address lines can be combined depending upon the way we want. what additional thing we have to have in the memory board is the control arrangement which will take care of the refresh control, which will take care of how to generate the read write signals, which will take care of the board select and chip select also of course comes from this decoder logics. So that part we will do in the next class. Which one?