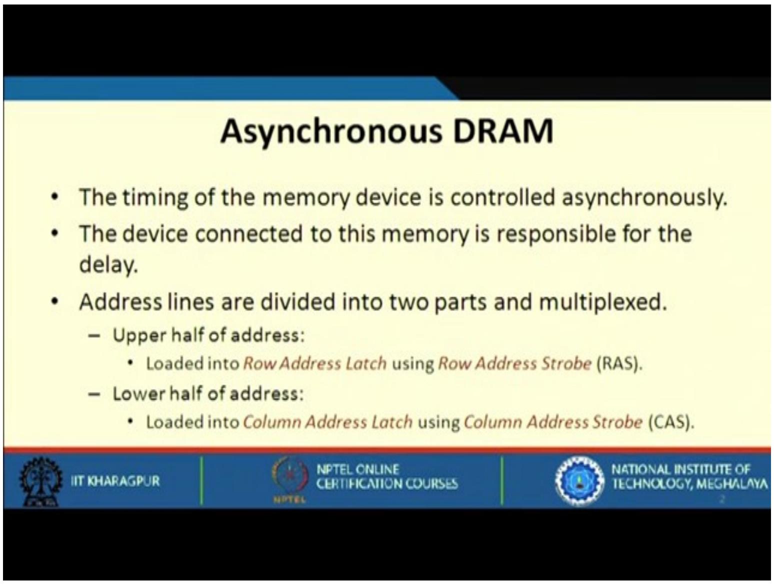
# Computer Architecture and Organization

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

## Lecture - 25
## Asynchronous DRAM

Welcome to lecture 25, where we will be discussing in detail about Asynchronous DRAM.

(Refer Slide Time: 00:33)



### Asynchronous DRAM

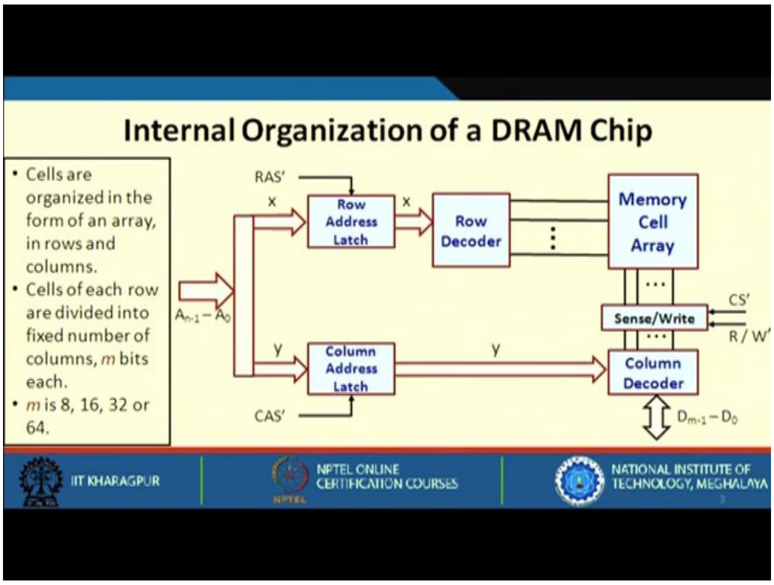
- The timing of the memory device is controlled asynchronously.
- The device connected to this memory is responsible for the delay.
- Address lines are divided into two parts and multiplexed.
  - Upper half of address:
    - Loaded into *Row Address Latch* using *Row Address Strobe* (RAS).
  - Lower half of address:
    - Loaded into *Column Address Latch* using *Column Address Strobe* (CAS).

IIT KHARAGPUR
NPTEL ONLINE CERTIFICATION COURSES
NATIONAL INSTITUTE OF TECHNOLOGY, MEGHALAYA

Here, the timing of the memory device is controlled asynchronously. So, the device connected to this memory is responsible for the delay. Ff you say that the processor is connected to the memory, then it is the responsibility of the processor to take care of that timing, when the data is available, etc. Here the address lines are divided into two parts and multiplexed. The upper half of the address is loaded into a latch called row address latch using a signal called row address strobe RAS.

And the lower half of the address is loaded into column address latch using column address strobe CAS. So, what do you mean by that it is multiplexed? That means, from outside we are providing a set of address, and that particular address is first hitting a particular row using this row address latch, and it is stored there by the signal RAS, and then the column address is provided. So, in the same address line the column address is provided and that is latched using CAS in column address latch. So, these are the two things how it is divided. We will be seeing this with examples later.

(Refer Slide Time: 02:24)

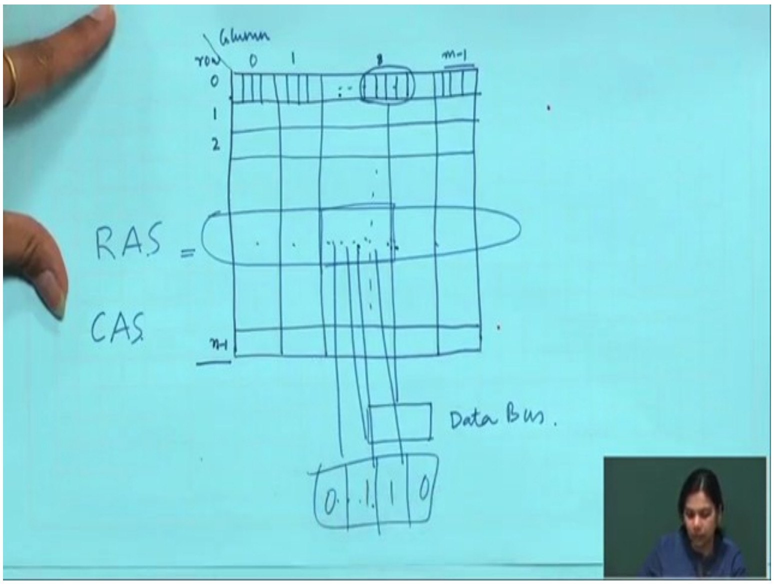


Internal Organization of a DRAM Chip

- Cells are organized in the form of an array, in rows and columns.
- Cells of each row are divided into fixed number of columns, *m* bits each.
- *m* is 8, 16, 32 or 64.

Let us now come to the internal organization of a DRAM chip. As I said the cells are organized in the form of an array in rows and columns. Cells of each row are divided into fixed number of columns of m bits each. Let me explain this.

(Refer Slide Time: 03:04)



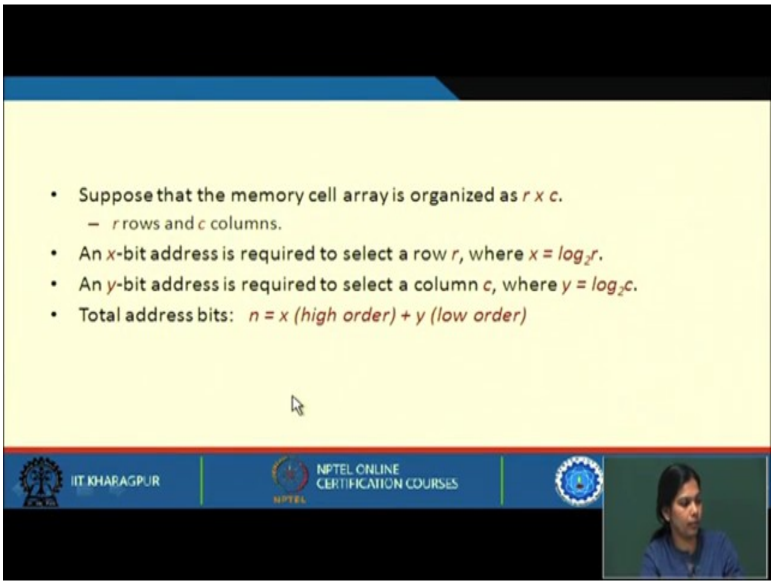
So, this is my memory. These are my rows. This is the first row, this is second row, and so on. So, the whole memory is divided into rows and columns. So, I have n rows and m columns. And in each of these m columns, I can have some bits. So, what I will do basically I will hit a particular row. And then there are m columns. I will hit a particular column.

Let us say this column. This is say 8th column I hit this and then, this particular column's data will be transferred to this sense or write circuit to the data bus. So, basically it is organized like this. It is divided into rows and columns, and each column is having some bits, let us say m. m is 8, 16, 32 or 64. Now let us see what we are doing here; we are applying the entire address 0 to n-1 --- let us say total address is n bits. So, you have your addresses 0 to n-1 that is divided into two parts, X and Y.

X address will go and hit to row address latch in response of the signal RAS. And this address is now fed to the row decoder, because there are number of rows. We need to select one particular row from set of these rows. The row decoder will be used to select one of the rows. Now this Y address will be provided to this column address latch in response of the signal CAS. And this Y will hit to the column decoder, because we also have many number of columns. And we do not know that in that particular row from which column I have to get the data; let say in this particular row I have to get the data from 10th column.

So, accordingly I have to pass a value here. So, a column decoder is in place that will decode a particular column that is available in this particular row. And so the sense or write circuit it will be read the data from that particular column, and it is transferred to the data bus lines. Or it can be written; in that case whatever data is in the data bus will be written into that particular column of that particular row. So, this is what is the internal organization of a DRAM chip.

(Refer Slide Time: 08:00)

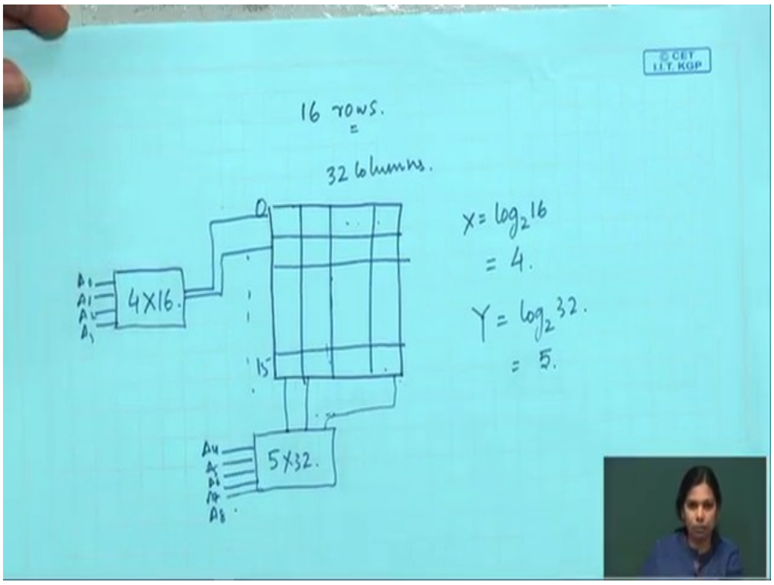


- Suppose that the memory cell array is organized as *r x c*.
  - *r* rows and *c* columns.
- An *x*-bit address is required to select a row *r*, where $x = log_2 r$.
- An *y*-bit address is required to select a column *c*, where $y = log_2 c$.
- Total address bits: *n = x (high order) + y (low order)*

So, suppose that a memory cell array is organized as row and column. So, r rows and c columns are there. An X bit address is required to select a row. The number of bits required for X will be log $_2$ r.

And now Y bit address is required to select a column, where Y will be log $_2$ c. So, the total address bit will be n that is X (high order bit) and Y (low order bit). I will just take a small example and explain this.

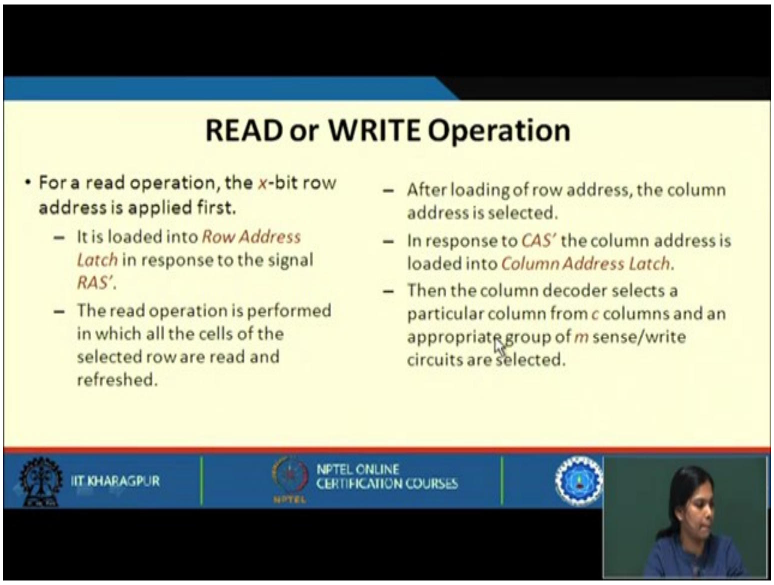
(Refer Slide Time: 08:53)



Here you see you have 16 rows So, how much size decoder you will require? You require a 4 x 16 decoder that will be connected. And accordingly each of these rows will be selected at a time. So, depending on this address a particular row will get selected. So, we have a total of 0 to 15, that is 16 rows. So, now, if you have 16 rows how many bits will be required; log $_2$ 16 because, r is the number of rows. So, X = 4.

Similarly, let us say you have 32 columns. If you have 32 columns that is Y, then you require 5 bits for your column address that will be provided. And you require a 5 x 32 decoder that will be connected to these. The 5 inputs will be let us say A4 A5 A6 A7 and A8 . So, if you have 16 rows and 32 columns you require a 4 x 16 decoder for the rows, and you need a 5 x 32 decoder to select a particular column.

(Refer Slide Time: 11:34)



## READ or WRITE Operation

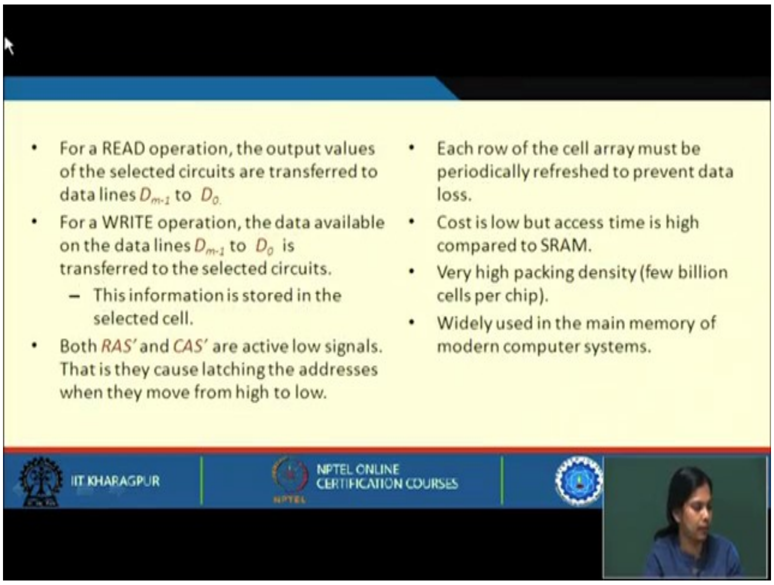
- For a read operation, the *x*-bit row address is applied first.
  - It is loaded into *Row Address Latch* in response to the signal *RAS'*.
  - The read operation is performed in which all the cells of the selected row are read and refreshed.
  - After loading of row address, the column address is selected.
  - In response to *CAS'* the column address is loaded into *Column Address Latch*.
  - Then the column decoder selects a particular column from *c* columns and an appropriate group of *m* sense/write circuits are selected.

Let us see how READ or WRITE operation will be performed. For a READ operation the X bit row address is applied first, because I have to go to a particular row first. And it is loaded into row address latch in response to the signal RAS. The READ operation is performed in which all the cells of the selected row are read and refreshed. After loading the row address the column address is selected. Now we have already selected a particular row. Now I have to know that within that row which column data to get. So, for doing, you have to apply the column address.

So, we apply the column address in the column address latch in response to CAS signal. Then the column decoder selects a particular column from c columns and an appropriate group of m sense or write circuits are selected. I will just take the same example here. Let us say I have selected this particular row. So, once we I select this particular row I put this address on RAS. And then I have to select any one of these columns. Then I apply CAS signal and let us say this particular column gets selected. And the m bits whatever bits there let us say 4 bit data will be output and will be available.

So, for reading what we need to do? We apply the column address first, we apply the row address first using RAS. Then we apply the column address using CAS. Then we select a particular column of that particular row and then finally, what we do we sense that set of bits from that particular column, and output it for a READ operation. And similarly for a WRITE operation whatever data is present in these 4 bits, let us say 0, 1, 1, 0; that set of data will be stored in this particular column of that particular row that has been provided.

(Refer Slide Time: 14:36)



- For a READ operation, the output values of the selected circuits are transferred to data lines $D_{m-1}$ to $D_0$
- For a WRITE operation, the data available on the data lines $D_{m-1}$ to $D_0$ is transferred to the selected circuits.
  - This information is stored in the selected cell.
- Both *RAS'* and *CAS'* are active low signals. That is they cause latching the addresses when they move from high to low.
- Each row of the cell array must be periodically refreshed to prevent data loss.
- Cost is low but access time is high compared to SRAM.
- Very high packing density (few billion cells per chip).
- Widely used in the main memory of modern computer systems.

This is how READ and WRITE operations take place. So, for a READ operation the output values of the selected circuit are transferred to the data line.

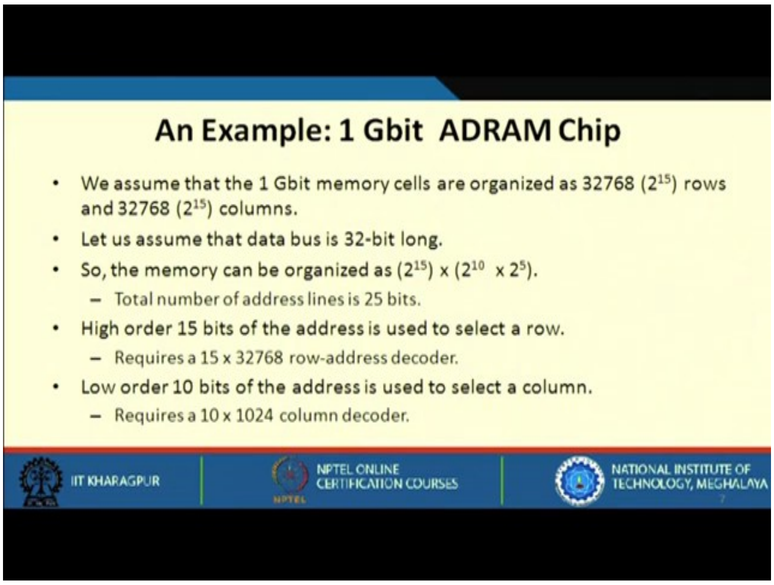
I said for a WRITE operation the data available on the data lines are transferred to the selected circuits. So, for read it is transferred to the data lines, for the write from data lines it will be stored in that particular step. This information is stored in that selected cell. Now both RAS’ and CAS’ are active low signals; this dash symbolizes this is active low. So, it will be high when it will be activated when the input is 0. So, they cause latching the addresses when they move from high to low. So, this will not happen when we move from low to high. Rather it will happen when we move from high to low. Each row of the cell array must be periodically refreshed to prevent data loss.

So, in this case each row of the cell array must be periodically refreshed. Cost is low because we have already seen that a single transistor and a single capacitor are required to build this. And of course, the cost is low, but the access time is high compared to SRAM, but billions of cells can be packed per chip.

And this particular kind of memory is used in main memory. We say that we have a main memory 4 GB. Can we extend it to 8 GB? Of course, you can extend it to 8 GB, provided you have that much available address bus.

Now I will take an example of 1 Gbit asynchronous DRAM chip.

(Refer Slide Time: 17:16)



## An Example: 1 Gbit ADRAM Chip

- We assume that the 1 Gbit memory cells are organized as 32768 ($2^{15}$) rows and 32768 ($2^{15}$) columns.
- Let us assume that data bus is 32-bit long.
- So, the memory can be organized as ($2^{15}$) x ($2^{10}$ x $2^5$).
  - Total number of address lines is 25 bits.
- High order 15 bits of the address is used to select a row.
  - Requires a 15 x 32768 row-address decoder.
- Low order 10 bits of the address is used to select a column.
  - Requires a 10 x 1024 column decoder.

IIT KHARAGPUR
NPTEL ONLINE CERTIFICATION COURSES
NATIONAL INSTITUTE OF TECHNOLOGY, MEGHALAYA

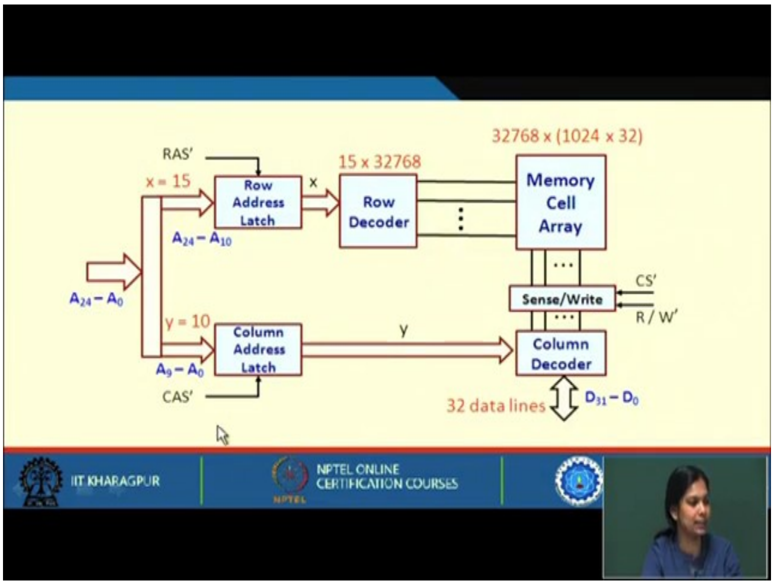
We assume that 1 GB memory cells are organized as 32768 that is, 2 to the power 15 rows and 32768 (2 to the power 15) columns. So, let us assume that the data bus is 32 bit long.

So, the memory cell can be organized as $2^{15}$ x $2^{10}$ x $2^5$. Total number of address lines is 25, because 15 bits we required to select a particular row, and 10 bit to select a particular column.

So, I will have $2^{10}$ columns and each of these columns will have $2^5$. So, total coming down $2^{15}$ x $2^{15}$ that is 1 gigabit. But total number of address line will be 15 for this row and, 10 bits for columns. So, 15 bits for selecting a row, and 10 bits for selecting a column, that is why 25 bits.

So, the high order 15 bits of the address will be used to select a row. So, we require a 15 x 32768 row address decoder. And the lower order 10 bits of the address will be used to select a column in that particular row. So, how many total columns are there? $2^{10}$, so we require a 10 x 1024 column address. So, total is 25, 15 for row and 10 for column.

(Refer Slide Time: 20:26)



32768 x (1024 x 32)

RAS'

x = 15

Row Address Latch

x

15 x 32768

Row Decoder

Memory Cell Array

$A_{24} - A_{10}$

$A_{24} - A_0$

Sense/Write

CS'

R / W'

y = 10

Column Address Latch

y

Column Decoder

$A_9 - A_0$

CAS'

32 data lines

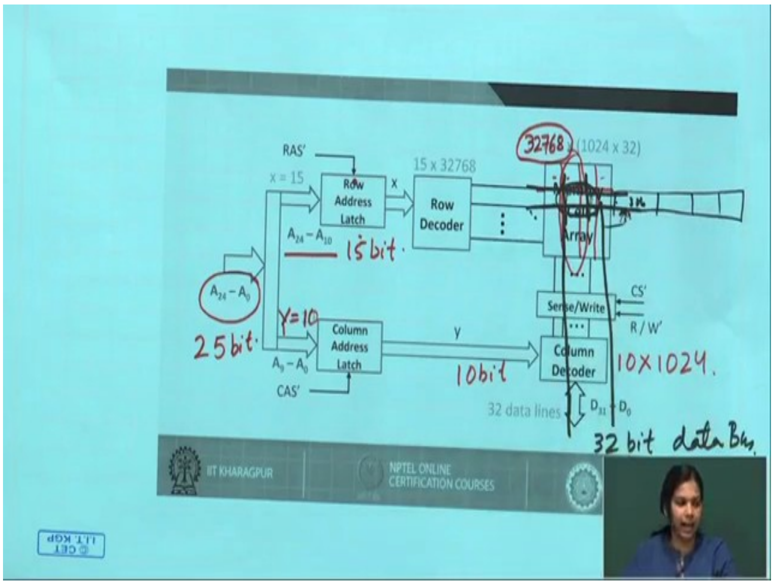
$D_{31} - D_0$

IIT KHARAGPUR

NPTEL ONLINE CERTIFICATION COURSES

Now, let us see how it is organized. The total address bit we are having is how much?

(Refer Slide Time: 20:33)



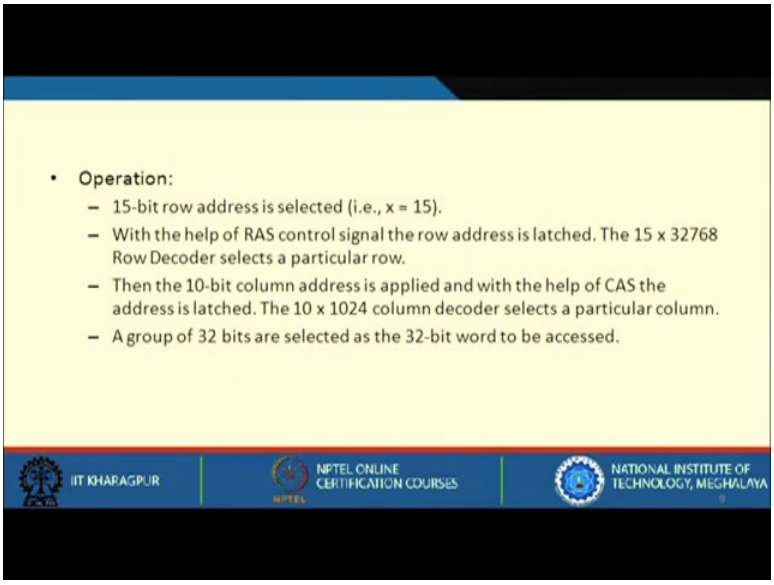
Total address bits are 25, 0 to 24. The high order 15 bits A10 to A24 will be fed to this row address latch in response to RAS signal. And then it is fed to the decoder that will decode that particular 15 bit address, and then it will map into one of the memory cells, one of the row of the total number of rows here is 32768 it will select any one row depending on this particular address. And then we apply the 10-bit column address; there is a column decoder. Depending on that, there are 1024 columns and any one of the columns will get selected and a 32-bit data will be transferred to the data bus.

So, let us see this from here, our total address is 25 bit. High order 15 bit is applied to row address latch in response to RAS signal. And then it is fed to the row decoder. So, these 15 bits fed to the row decoder will select any one of these 32768 rows. And any one of the row will get selected. Now I have 10-bit column address; so these 10 bit columns address A0 to A9. Low order 10-bit column address will be applied to this column address latch in response to the signal called CAS. This will be fed to the column decoder.

Now, this column decoder is a 10 x 1024 column decoder. It accepts 10 bits and it will select any one of the 1024 columns. Let us say it has selected this particular col column. I have selected this particular row. And now this column decoder has selected this particular column. Now, on this row this particular column got selected, and this will be transferred to your data bus.

So, this is your 32 bit data bus, where this particular data will be transferred. So, how we are doing this decoding? We apply the total 25-bit address. High order bit goes to the row address. The row decoder will select a particular row. Then the low order bits are applied to the column address. This is 10 bit through this column decoder any one of the 1024 columns will get selected. So, we have a row, we have a column and then 32 bits of data transferred. So, this is how it happens in asynchronous DRAM.

(Refer Slide Time: 25:24)

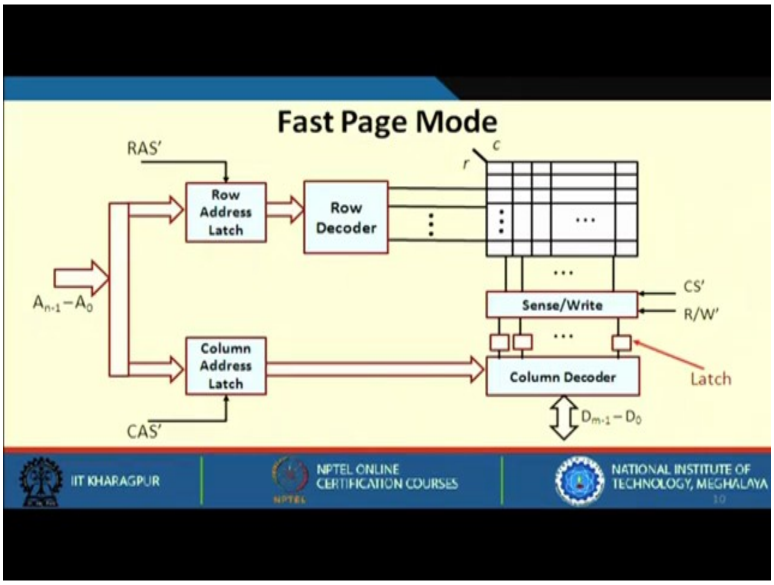


- Operation:
  - 15-bit row address is selected (i.e., x = 15).
  - With the help of RAS control signal the row address is latched. The 15 x 32768 Row Decoder selects a particular row.
  - Then the 10-bit column address is applied and with the help of CAS the address is latched. The 10 x 1024 column decoder selects a particular column.
  - A group of 32 bits are selected as the 32-bit word to be accessed.

IIT KHARAGPUR
NPTEL ONLINE CERTIFICATION COURSES
NATIONAL INSTITUTE OF TECHNOLOGY, MEGHALAYA
9

Now, this is what just now I explained. 15-bit row address is selected, with the help of RAS control signal. The 15 x 32768 row decoder selects a particular row. And then the 10-bit column decoder is applied. And with the help of CAS the address is latched. The 10 x 1024 column decoder will now select a particular column. And then a group of 32 bits may be accessed. Next we will discuss another thing that is fast page mode access.

(Refer Slide Time: 26:16)



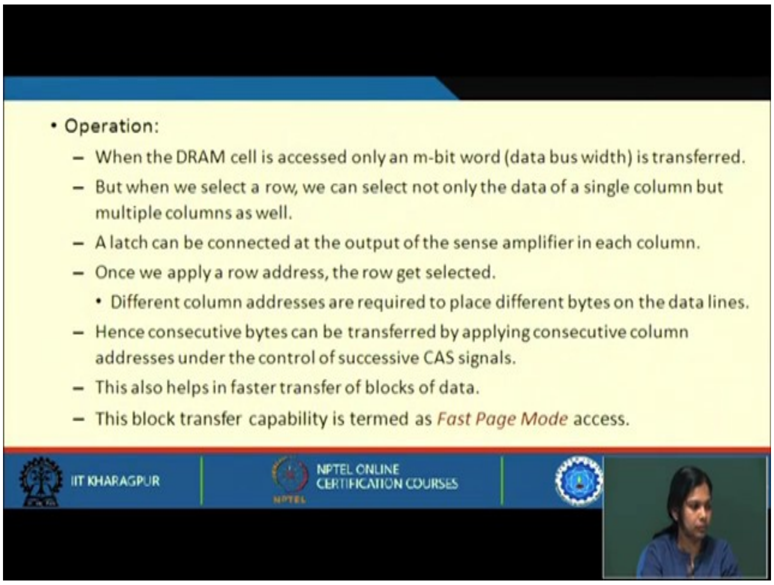
Fast Page Mode

If you recall this what we were doing, when we are applying this row address and column address, we are in this particular row. And in this particular row there are any columns, but what we are doing? We are actually bringing a data of only one column, but accessing we are doing of this particular row, but we are only accessing one column data. What if you want to access other set other columns data as well in the same group? Because I have already said that we never bring a single bit of a single word rather we bring multiple words that is a block, we transfer a block. So, in this case you can see that this particular column is selected now.

What if there is a column counter and then I can also select other values of the same column? So, there are other columns, and all other columns are having this 32-bit data. Instead of again selecting a row, we have already selected a row. We have already selected a column. And let us now select next set of columns. So, then the access will be much faster. Because I have another counter that will count that will bring me to the next column address, and I will fetch the next data, and so on. This particular feature is known as fast page mode, provided the row address through this row decoder, provided the column address to this column decoder.

But only we need to have a latch in this place, because output of this will be stored here and then it will be transferred to the data lines. Then the next column address will be stored and then it will be transferred, then the next column, then the next column and so on. So, in a single row access what we are achieving, we are not only accessing one bit of data, but one set of data that is 32 bit in this case.

(Refer Slide Time: 29:20)



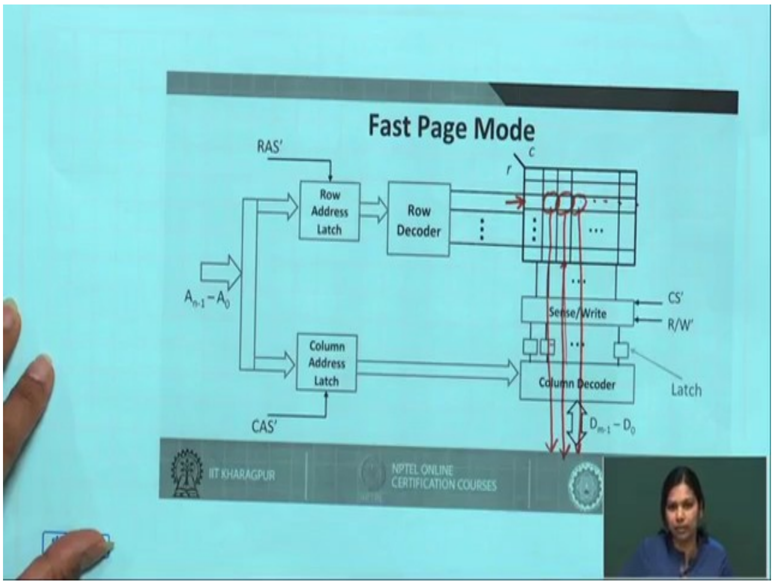
- Operation:
  - When the DRAM cell is accessed only an m-bit word (data bus width) is transferred.
  - But when we select a row, we can select not only the data of a single column but multiple columns as well.
  - A latch can be connected at the output of the sense amplifier in each column.
  - Once we apply a row address, the row get selected.
    - Different column addresses are required to place different bytes on the data lines.
  - Hence consecutive bytes can be transferred by applying consecutive column addresses under the control of successive CAS signals.
  - This also helps in faster transfer of blocks of data.
  - This block transfer capability is termed as *Fast Page Mode* access.

So, just now what I have said, when the DRAM cell is accessed only m bit of word that is width of the data bus is transferred. But when we select a particular row, we can select not only the data of a single column, but multiple columns as well. Just now I have mentioned this. We can simply do this by connecting a latch to the output of the sense amplifier in each column.

Once we apply a row address the row gets selected. And then different column addresses are required to place different bytes on the data line; that means, we already accessed a particular row, now I have already access to one particular column; now to access the next column, I just have to give the next column address, next column address, and so on. So, I need not have to select the row again. Hence consecutive bytes can be transferred by applying consecutive column addresses under the control of successive CAS signals. But we need to give successive CAS signals for that. This also helps in faster transfer of blocks of data because this is useful when we transfer blocks of data.

So, this block transfer capability is termed as fast page mode access that I just now explained.

(Refer Slide Time: 31:20)



Fast Page Mode

Here you can see that I hit this particular row, I hit this particular column, and then the data is available. Now I apply the next column address through this column decoder and I get the next data, then the next, then the next, and so on. We need not have to apply this row address again. So, this makes the transfer much faster. This is termed as block transfer capability, which is fast page mode access.

So, we came to the end of lecture 25. In the next lecture we will see synchronous DRAM. Here we have discussed about asynchronous DRAM, where the timing is dependent on whichever model is connecting to the memory. But in synchronous DRAM we will see that there is a clock involved.

Thank you.