

Register File Design and Memory Design

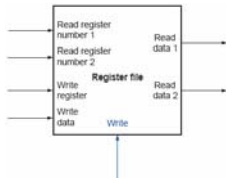
Dr. Arjan Durrresi
Louisiana State University
Baton Rouge, LA 70810
Durrresi@Csc.LSU.Edu

These slides are available at:
http://www.csc.lsu.edu/~durrresi/CSC3501_05/



- Register File
- Memory
- SRAM
- DRAM
- Memory Chip

Register File

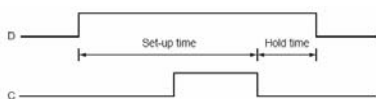


- The register file includes 32 32-bit registers, as it is needed for 32 general purpose registers of MIPS architecture
- This register file makes possible to simultaneously read from two registers and write into one register as it is appropriate for MIPS processor.

Register File Functioning

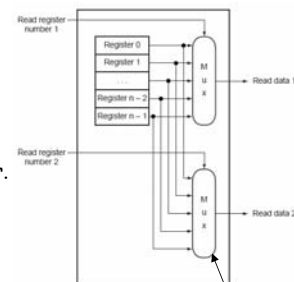
- The given register file functions as follows:
 - any value provided on 5-line Read register number 1 port makes that the content of corresponding register is provided on 32-line Read data 1 port
 - any value provided on 5-line Read register number 2 port makes that the content of corresponding register is provided on 32-line Read data 2 port
 - on the falling edge of write line, values that appear on 32-bit Write data port are written into the register with the number on 5-line Write register port.
- Note that requirements for set-up time (and hold time) also apply here.

Set-up and hold time requirements for a D flip-flop with a falling-edge trigger



- The input must be stable a period of time before the clock edge, as well as after the clock edge.

Register File Design: Read Part

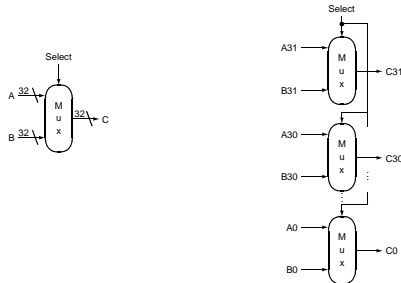


- This is a design at a level of register and complex multiplexer.

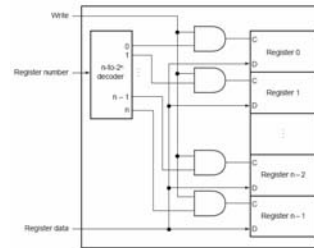
Do you understand? What is the "Mux" above?

Abstraction

- Make sure you understand the abstractions
- Sometimes it is easy to think you do, when you don't



Register File Design: Write Part



- This is a design at a level of register and decoder.

Introduction to Memory Design

- Main memory is built in one of two technologies:
 - SRAM- Static Random Access Memory
 - DRAM- Dynamic Random Access Memory
- A memory is normally built using a number of memory chips.
- Memory chips have specific configurations given as a product of two number, e.g.
 - 128K*1 - 128K addressable locations with 1 bit in each location, i.e. width of read/write operations is 1 bit
 - 16K*8 - 16K addressable locations with 8 bits in each location, i.e. width of read/write operations is 8 bits
- Notice that two chips above accommodate identical number of bits (128K bits).
- Both memories are *volatile*.

SRAM and DRAM: 1 Bit Memory Cell

- In SRAM technology, three-state D-latch is a basic building block, i.e. basic memory cell.
 - Internally, D-latch can have a state corresponding to 0 or 1.
- In DRAM technology, a basic memory cell is build around one capacitor coupled with one transistor. The value in the cell is stored as a charge. A charge can not be stored indefinitely and DRAM chips must be periodically refreshed.
 - Since charge can be kept for several msec, 1-2% of time is used for refreshing.

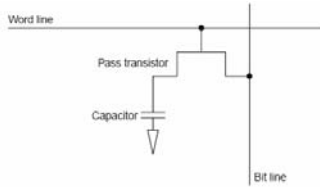
SRAM Technology Characteristics

- SRAM - Static Random Access Memory technology is normally used for caches
- In comparable technologies, SRAM cycle time is about 8 to 16 times faster than DRAM, e.g. currently 0.5-5 nsec.
- But, SRAM chip capacity (as well as density) is roughly 4 to 8 times less than that of DRAM
- Also SRAM is more expensive, e.g. 1GB in 2004 \$4,000 - \$10,000 for SRAM and \$100 - \$200 for DRAM
- In addition, SRAM chips have higher power consumption and power dissipation than DRAM chips
- Thus SRAM designs are concerned with speed, while in DRAM designs the emphasis is on cost per bit and capacity.

DRAM Technology Characteristics

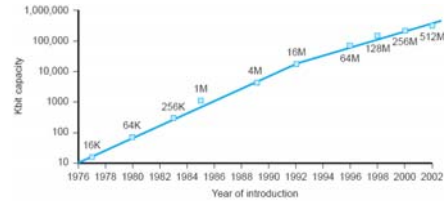
- Since 1975, the main memory is composed of semiconductor DRAM's (Dynamic Random Access Memory).
- DRAM chip capacity had been growing at rate of about 4 times every three years, while lately growth slowed down to 2 times every two years.
- Currently, maximum DRAM chip capacity is 512M bits with an access time in the range 40-60 nsec and a cycle time of about 80 nsec.
- Access time & cycle time are two measures of memory latency:
 - access time - the time between a read is requested and when the desired content arrives,
 - cycle time - the minimum time between two memory requests.
- For DRAM technology cycle time is longer than access time.
- For SRAM technology access time and cycle time are identical.

DRAM

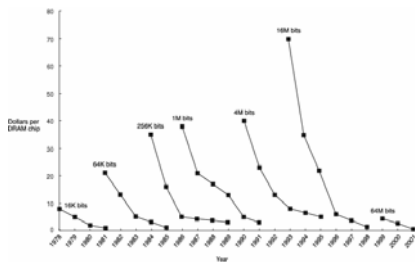


- A single-transistor DRAM cell contains a capacitor that stores the cell contents and a transistor used to access the cell.

Growth of Capacity per DRAM chip

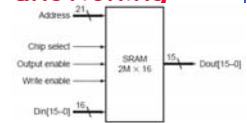


Prices of Six Generations of DRAMs



Memory Chip Functioning

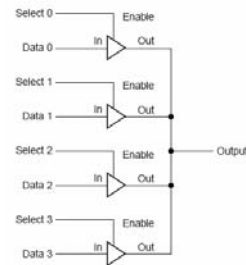
- Example: 32K×8 chip
 - read and write operations are 8 bits wide
 - there are 32K addressable locations
- CS (Chip select) has to be set for either reading or writing
- R (Read enable)=0 & W (Write enable)=0 → chip is not being accessed
- R=0 and W=1 → write values at Din lines into the chip address at Address lines
- R=1 and W=0 → read into Dout lines values from the chip address at Address lines
- R=1 and W=1 → not allowed
- Two designs of memory chip:
 - Basic structure design
 - Typical organization design



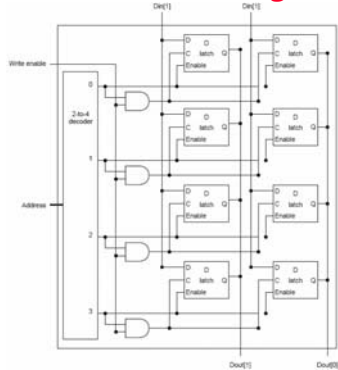
Design of Memory Chip Basic Structure

- The design of the basic structure of SRAM chip uses some ideas from the register file design e.g. the write parts in two designs are identical. The main differences are in read part design. In the memory chip with the usage of three-state D-latches a multiplexer is eliminated. E.g. for 32K×8 SRAM chip, a multiplexer with 32K inputs each input having 8 lines would be needed.
- But for design of the basic structure of SRAM chip, we still need a very large decoder. E.g. for 32K×8 SRAM chip, a decoder with 15 input lines (that is not so bad) and 32K output lines (that is bad) is required. Typical organization of SRAM uses two level decoding that eliminates need for that very large decoder.

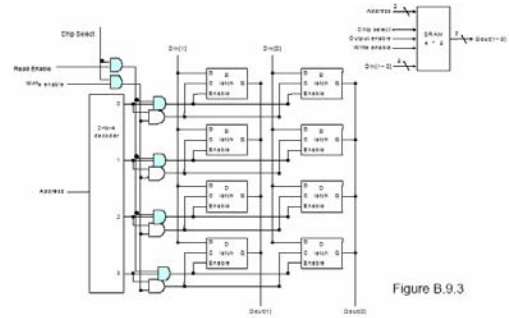
Three-state buffers are used to form a multiplexor.



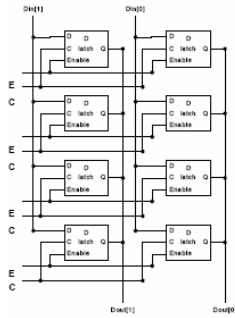
Basic Structure Design of 4x2



Basic Structure Design of 4x2 SRAM Chip

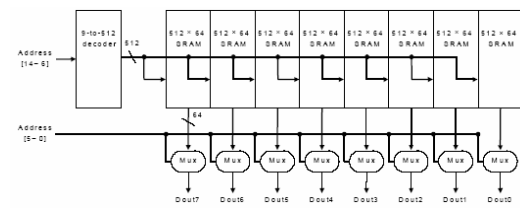


4x2 Array of D- Latches

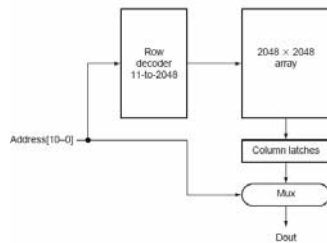


Typical Organization Design

- Example: Design (read part only) a typical organization (i.e. two level decoding design) of 32Kx8 SRAM chip that uses 512x64 arrays of D-latches.
- Note: The used arrays have to have a bit capacity equal to a number of addressable locations in the chip, e.g. in this example that condition is satisfied since $512 \times 64 = 32K$. A number of arrays used should be equal to the number of bits in each memory location.



DRAM Organization



- The row access uses 11 bits to select a row, which is then latched in 2048 1-bit latches. A multiplexor chooses the output bit from these 2048 latches.
- The RAS and CAS signals control whether the address lines are sent to the row decoder or column multiplexor.

Main Memory Specification

- A memory has identical inputs and outputs as memory chips, except that CS does not exist. But the specification of a memory should include:
 - a. memory capacity (usually in bytes),
 - b. memory addressability, i.e. smallest unit that has its address,
 - c. width of read/write operations, i.e. a number of bits that can be read or written from/to memory.
- Operations on memory: reading from memory and writing into memory;
 - RE=0 and WE=0 → memory is not being accesses
 - RE=0 and WE=1 → writing into memory
 - RE=1 and WE=0 → reading from memory
 - RE=1 and WE=1 → not allowed

Error Correction

- **Error-detecting code** A code that enables the detection of an error in data, but not the precise location, and hence correction of the error.
 - A 1-bit parity scheme
- **Error-correcting codes** that will detect and allow correction of an error.

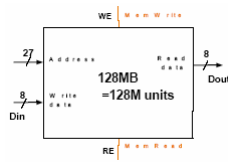
Error Correction

- A 1-bit parity code is a *distance-2 code*, which means that if we look at the data plus the parity bit, no 1-bit change is sufficient to generate another legal combination of the data plus parity.
- To detect more than one error or correct an error, we need a *distance-3 code*, which has the property that any legal combination of the bits in the error correction code and the data have at least 3 bits differing from any other combination.

Data	Code bits	Data	Code bits
0000	000	1000	111
0001	011	1001	100
0010	101	1010	010
0011	110	1011	001
0100	110	1100	001
0101	101	1101	010
0110	011	1110	100
0111	000	1111	111

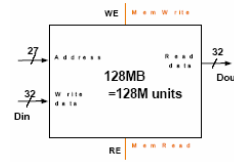
Main Memory Specification: Example 1

- Provide inputs and outputs of 128MByte memory with 8-bit read/write operations and byte addressability.
- Note that 8-bit read/write operations requires byte addressability.



Main Memory Specification: Example 2

- Provide inputs and outputs of 128MByte memory with 32-bit read/write operations and byte addressability.



Main Memory Specification: Example 3

- Provide inputs and outputs of 128MByte memory with 32-bit read/write operations and 32-bit addressability.



Summary



- Register File
- Memory
- SRAM
- DRAM
- Memory Chip