

# Cache Simulator: Block Set Associative and Least Recently Used

Theoni Anne N. Lim

De La Salle University, theoni\_lim@dlsu.edu.ph

Vincent Alvin S. Marquez

De La Salle University, vincent\_marquez@dlsu.edu.ph

Joaquin Lorenzo T. Sitoy

De La Salle University, joaquin\_sitoy@dlsu.edu.ph

Stanley Yale Z. Zeng

De La Salle University, stanley\_yale\_zeng@dlsu.edu.ph

## Links:

Github: [https://github.com/stanstan50/Cache\\_Simulation.git](https://github.com/stanstan50/Cache_Simulation.git)

Website: <https://cache-simulation-jmwww.onrender.com>

Video Demo: [https://drive.google.com/file/d/1TII1ePo6A03tIWleCH0q6T8NpMwVLgr\\_/view?usp=sharing](https://drive.google.com/file/d/1TII1ePo6A03tIWleCH0q6T8NpMwVLgr_/view?usp=sharing)

## I. INTRODUCTION

The Cache Simulator demonstrates how a block set associative cache mapping using the Least Recently Used algorithm can be used to store a sequence of data. The web application uses the node.js framework and was deployed using Render.

The front end was coded using standard HTML, CSS, and Javascript for user interaction. Inputs are manually typed in, but incrementors are also available. Dropdown selection boxes are available to switch between a block input (integers) or an address/word input (hex). A sequence of space-separated blocks or words are to be inputted in the Program Flow. These represent the data from main memory to be accessed sequentially. Pressing the “Simulate” button will output the cache snapshot to the right and the “Export As Text File” will export a text file of the time calculations and corresponding inputs.

The back end has a main simulate() function that initiates the cache accessing simulation. Functions like getUserInput() formats the inputs from the HTML-side into something more digestible for other functions while displayCache() feeds the cache snapshot (represented as a 2D array) back to the front end for display. Many helper functions such as parseProgramFlow(), hexadeciToDeci(), and createSetRow() are used for formatting, data conversion, and data creation. Inputs are error-checked before sending to the simulate() function.

## II. TEST CASES AND RESULTS

Inputs are shown on the left while the output/cache snapshot is shown on the right. Below the input box are the access time calculations. Indices for each test case are to the left of the screenshot column.

Table 1: Valid-Normal Input/Output

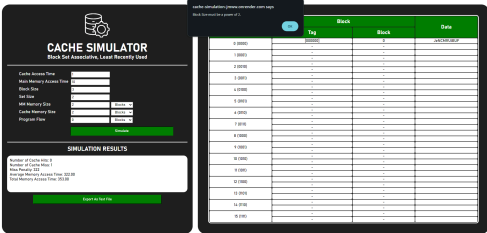
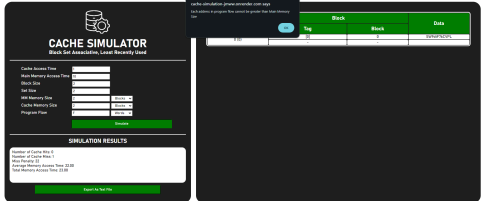
0		9	
---	--	---	--

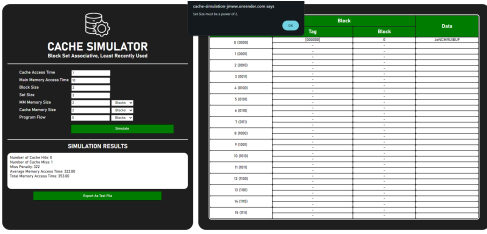
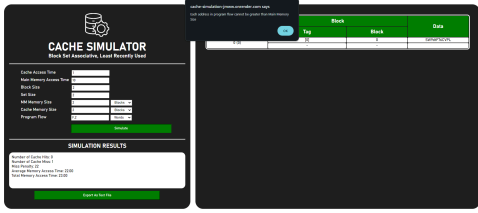
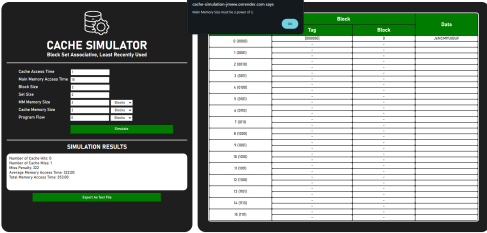
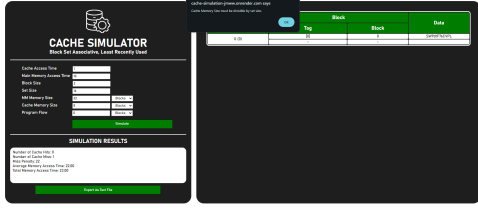
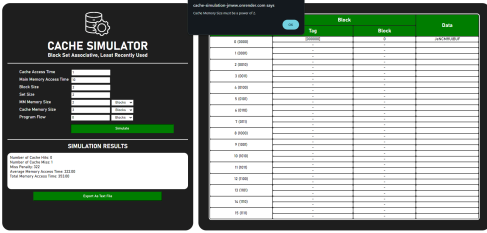
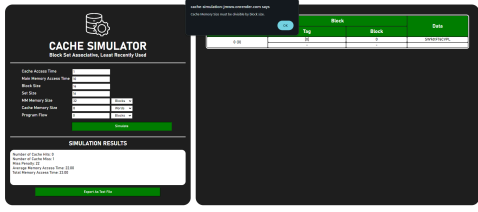
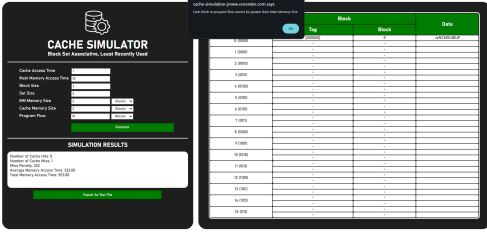
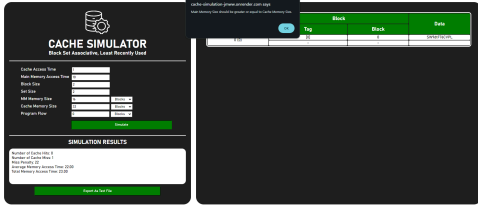
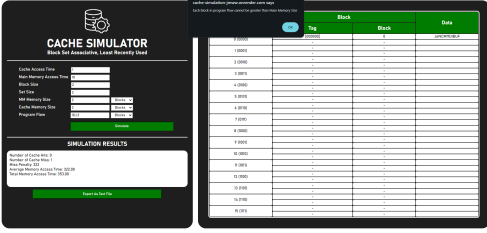
1		10	
2		11	
3		12	
4		13	
5		14	
6		15	

Table 2: Valid-Special Input/Output

0	
1	
2	

Table 3: Invalid Input

0		6	
---	---	---	--

1		7	
2		8	
3		9	
4		10	
5			

### III. ANALYSIS AND DISCUSSION

To verify the results of the simulator, manual computations were written to clarify the specifications of the block set associative, LRU mapping.

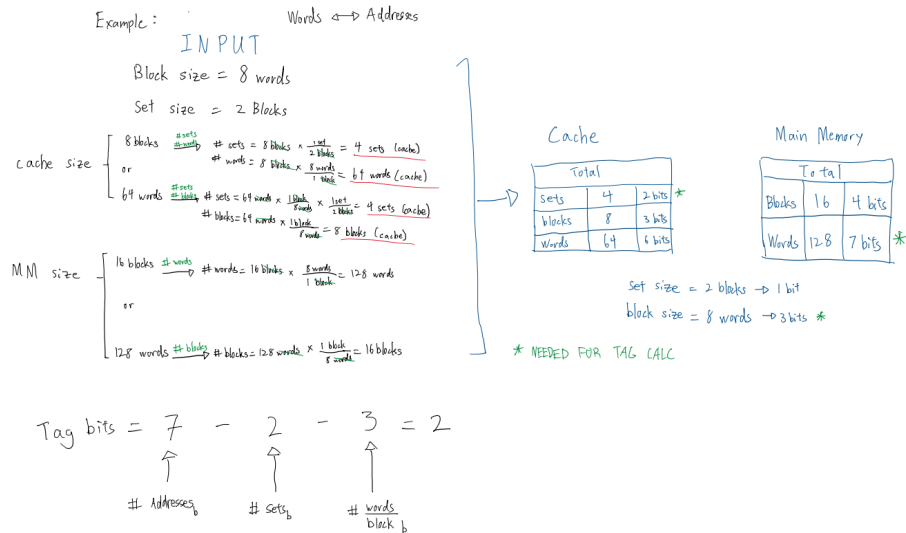


Figure 1: Manual Computation of Test Case 1 of Valid/Normal Input

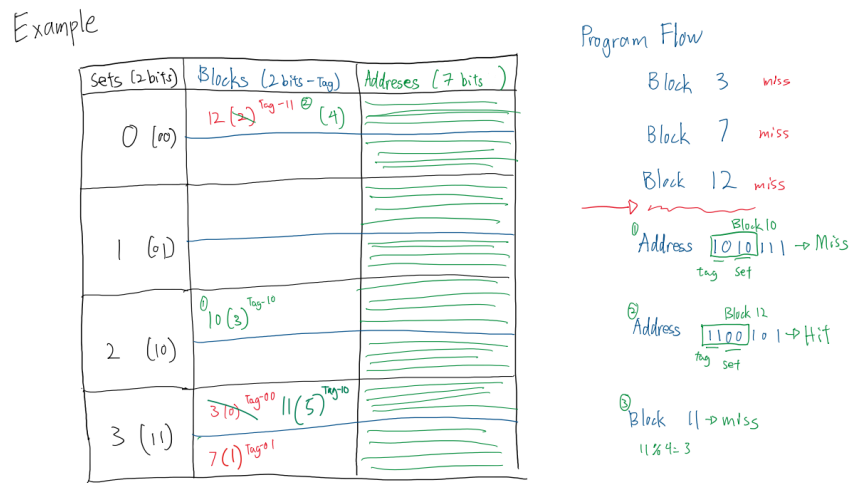


Figure 2: Manual Computation of Test Case 1 of Valid/Normal Input (cont.)

The figures above show hand calculations of Test Case 1 from Table 1 and in the video demonstration, this was the second test case to be shown. As seen from the example, blocks 3, 7, 12, and 10 were misses, resulting in them being stored in their respective sets using the modulo operation, with the age in parentheses. The second block 12 initiated a hit, updating the age in Set 0, Block 0. Block 11 resulted in a miss. The modulo operation allocates it to Set 3, but since it's already full, the least recently used value (block 3), gets replaced. This manual computation reflects the snapshot shown in the simulator. For a detailed analysis of the process, the main program flow is as follows:

```
function simulate() {
  console.log("=====");
  console.log("Simulating...");
  console.log("=====");
  getUserInput();
  simulateCache();
  displayCache();
  displayResults();
}
```

Figure 3: Main Program Flow

```

blockSize = $("input[name='blocksize']").val();
setSize = $("input[name='setsize']").val();
MMSize = $("input[name='mmsize']").val();
MMType = $("select[name='mmtype']").val();
cacheSize = $("input[name='cachesize']").val();
cacheType = $("select[name='cachetype']").val();
programFlow = $("input[name='progflow']").val();
programFlowType = $("select[name='progflowtype']").val();
cacheAccessTime = parseFloat($("input[name='cacheAccessTime']").val());
MMAccessTime = parseFloat($("input[name='mmAccessTime']").val());
wordsPerBlock = blockSize;
blocksPerSet = setSize;
originalCache = cacheSize;
originalMM = MMSize;

```

```

function parseProgramFlow(programFlow) {
    //blocks (block number) in decimal
    if(programFlowType == "blocks") {
        //remove any non-digit characters before and after the string
        programFlow = programFlow.replace(/^\D+|$/g, '');

        //trim
        programFlow = programFlow.trim();

        //splits the string into an array of any non-digit characters
        programFlowArr = programFlow.split(/\D+/);

        //converts the array of strings into an array of numbers
        programFlowArr = programFlowArr.map(Number);
    } else { //words (addresses) in hex, considers A-F hex values
        //remove any non-digit characters before and after the string
        programFlow = programFlow.replace(/[^0-9A-Fa-f]+/g, '');
    }
}

```

Figure 4: Code Snippets of getUserInput() (left) and parseProgramFlow() (right)

The function, getUserInput(), uses jquery to extract the text from the website's input fields. The function, parseProgramFlow() is then called to parse the input from the program flow field, using any unrelated characters such as special characters or non-hexadecimal letters as separators between the actual data. If no valid blocks or hexadecimal characters are present, the program flow will default to 0 as an input. This ensures that the program doesn't crash despite invalid inputs.

```

function simulateCache(){
    if(isCacheSizeWord == true){
        cacheSize = cacheSize/wordsPerBlock;
    }
    if(isMMSizeBlock == true){
        MMSize = wordsPerBlock * MMSize;
    }
    let cache = initializeCacheMemory(cacheSize, blocksPerSet);
    let MMbits = getMMBitsWords(MMSize);
    let numOfSets = cacheSize/blocksPerSet;
    let setBits = getSetBits(numOfSets);

    for(let ctr = 0; ctr < programFlow.length; ctr++){
        let current = programFlow[ctr];
        let tagBits = getTagBits(current, wordsPerBlock, setBits, MMbits, isProgramFlowBlock);
        if(isProgramFlowBlock == true){
            setNum = getSetNumBlock(current, numOfSets);
        } else {
            setNum = getSetNumWord(current, numOfSets, setBits, MMbits);
        }
        let result = searchSet(tagBits, setNum, cache);

        if(result == 1){ //if cache hit
            cache = updateFoundBlock(tagBits, setNum, cache);
            hitCount++; //increment hit count
        } else { //if cache miss
            cache = insertMissingBlock(tagBits, setNum, cache, setBits);
            missCount++; //increment miss count
        }
    }
    console.log(cache);
    cacheSnapshot = cache;
}

```

Figure 5: Code Snippet of simulateCache()

Function simulateCache() is then called which performs the main operations and computations to simulate cache accessing and obtain the final cache snapshot. The function first checks whether the cache and main memory are in either blocks or words. It then converts the Cache Memory Size into cacheSize (cache size in blocks) and MM Memory Size into MMSize (main memory size in words) for the simplicity of using a single variable. Variable cacheSize will be used for creating the 2D array representation of the cache memory while MMSize will be used to compute the number of bits needed to represent an address. A for-loop was implemented to iterate over each element in the formatted program flow array. For each element, the program extracts the bits needed to search for an address in the cache, checking for a cache hit or miss. Should a hit occur, the program updates the block's age for the LRU algorithm and increments the hit counter. For a miss, the program inserts the missing block into the cache and similarly updates the age and miss counter. The final representation is saved in the cacheSnapshot variable. This whole process is to simulate, step-by-step, the cache operations in order to ensure the program is accurate to real world situations.

```

function displayCache() {
    document.querySelector('.right-side').innerHTML = createTable();
}

```

Figure 6: Code Snippet of displayCache()

```

function displayResults() { // calculations in nanoseconds
    totalCount = hitCount + missCount;
    missPenalty = cacheAccessTime + blockSize * MMAccessTime + cacheAccessTime; // 1 cache check + blockSize * memory access + 1 cache read
    averageMemoryAccessTime = (hitCount/totalCount) * cacheAccessTime + (missCount/totalCount) * missPenalty; // hits/total * cacheAccessTime + misses/total * missPenalty
    totalMemoryAccessTime = (hitCount * blockSize * cacheAccessTime) + (missCount * blockSize * (MMAccessTime + cacheAccessTime)) + (missCount * cacheAccessTime);

    //fix to 2 decimal places
    averageMemoryAccessTime = averageMemoryAccessTime.toFixed(2);
    totalMemoryAccessTime = totalMemoryAccessTime.toFixed(2);
    missPenalty = missPenalty;

    document.querySelector('.simulresults-field').innerHTML = `
    <p>
        Number of Cache Hits: ${hitCount} <br>
        Number of Cache Miss: ${missCount} <br>
        Miss Penalty: ${missPenalty} <br>
        Average Memory Access Time: ${averageMemoryAccessTime} <br>
        Total Memory Access Time: ${totalMemoryAccessTime} <br>
    </p>
    `;
    document.getElementById('exportres').style.display = 'block';
}

```

**Figure 7: Code Snippet of displayResults()**

Function displayCache() is then called to update the final snapshot on the right hand side of the front page. Function displayResults() computes the memory access time, taking into account the hit count and miss penalty. The formulas are as shown in the figure above. The final product of displayResults is shown at the bottom of the front page, below the inputs, and can be downloaded as a text file.

#### IV. CONCLUSION

When accessing data, it is common to have data that is more frequently used than others. Thus, it is efficient to have storage devices such as caches closer to the processor for faster access. Different cache mapping algorithms are used to determine how to store the data and simulators are useful for visualizing the resulting cache structure. Compared to other mappings such as direct mapping or fully associative, block set associative provides a balance between quick access and data overlap prevention. Test cases in the documentation show how the simulator can be used, illustrating the type of input needed and corresponding output, including error-handling. This project provided valuable insight on the inner workings of cache memory.