

University of Pennsylvania  
Department of Electrical and System Engineering  
System-on-a-Chip Architecture

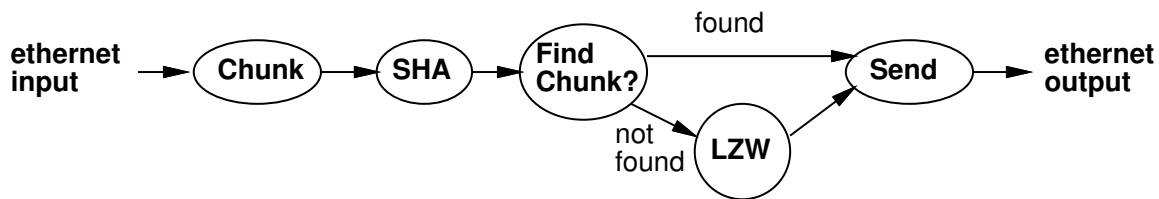
ESE532, Fall 2018 Deduplication and Compression Project Wednesday, October 24

---

**Due:** Friday, December 7, 5:00PM

## 1 Goal

Develop a compressor that can receive data in real time at modern ethernet speeds (1Gb/s, maybe 10Gb/s) and compress it into memory using deduplication and compression. Specifically, we'll look at Content-Defined Chunking to break the input into chunks, SHA-256 hashes to screen for duplicate chunks, and LZW compression to compress non-duplicate chunks. Your goal for implementation is to achieve real-time guaranteed support of 1Gb/s input stream, but you may need to consider intermediate goals (e.g. 100Mb/s, 300Mb/s) along the way. In any case, you will sketch the solution to scale to 10Gb/s even if you cannot implement it.



This is a 5 week project assignment; the intent is to allow you to plan and execute a significant, open-ended design exploration and mapping. You will not achieve the implementation goal or the course learning goals by trying to do this in one week. We give you milestones to help provide some structure, but the milestones are minimal and *doing the minimum to hit the milestone each week will be insufficient* to get you where you need to be at the end. We are giving you flexibility in planning and ordering rather than lock-step specifying exactly what you need to do each week.

- Project work is done in teams of 3. You select partners during first week.
- Collaboration between teams is limited as specified on the [course web page](#).
- Individual milestone and final writeups, including individual solution of exercises.
- The spirit of this exercise is to optimize the SoC mapping of the algorithm. As such, explorations of alternate solutions that change the algorithms and generally optimize the solution for hardware and software are out of scope. Explorations that tweak or tune the algorithms slightly to better exploit the SoC hardware are potentially in scope.

## 2 Final Report

- Describe your single ARM processor mapped design. [1 page]
  - Key parameters in the solution
  - Performance achieved and energy required
  - Characterization and breakdown of time spent in the major components
- Describe your final Zynq XC7Z020 mapped design. [5 pages]
  - Performance achieved and energy required
  - Key design aspects: task decomposition, parallelism, mapping to Zynq resources, include diagrams to support
  - Be clear where each component of the final design is performed (e.g., ARM, NEON vector, FPGA logic).
  - Model to explain performance, area, and energy of design
  - Current bottleneck preventing higher performance
- Describe design that achieves 10Gb/s and estimate chip area and costs. [3 pages]
  - Key design aspects. Build on description in first section, describing differences for this design from the Zynq XC7Z020 design.
  - Use or expand model introduced in first section to show how you meet the 10Gb/s encoding problem.
  - Custom implementation area estimate with area model support
- Describe how you validated your implementations, including the real-time guarantee for the input rate. The real-time validation likely includes both arguments about the way the code is written and mapped to the Zynq and your testing methodology. [2 pages]
- Describe the key lessons you learned from this design experience. [1 page]
- Describe design space explored and show graphs and models to support design selection. [any number of pages as needed]
- Describe who did what. [1 page]
- Submit a tar file with your final design; there will be a separate project assignment link for your code submission.

### 3 Milestones

We will provide precise requirements for milestones each week. These may include a few exercises to help prepare you for questions that may be on the final in addition to the project specific components. Milestones and feedback feed into the final report. In most cases, the milestones can serve as a first draft of a component of your report, and the feedback we give you will help provide guidance on how to refine it for the report.

- Analysis, parallelism, and teaming [11/2]
- Functional version and design space [11/9]
- Input integration and energy estimate [11/16]  
(intervening lectures and this milestone specification will clarify what we want for energy estimate)
- 200Mb/s design and custom area estimates for that design [11/30]  
(intervening lectures and this milestone specifications will specify area model)
- Final Report [12/7]

## 4 Components

The components we will use are standard enough that the wikipedia pages are useful, and there are several other nice tutorial blog posts out there. Here's a roundup of starting points.

- Content-Defined Chunking (Rabin Fingerprint)
  - <https://moinakg.wordpress.com/tag/rabin-fingerprint/>
  - <https://restic.net/blog/2015-09-12/restic-foundation1-cdc>
  - [https://en.wikipedia.org/wiki/Rabin\\_fingerprint](https://en.wikipedia.org/wiki/Rabin_fingerprint)
- SHA 256 Hashing
  - <https://tools.ietf.org/html/rfc6234>
  - <https://en.wikipedia.org/wiki/SHA-2>
- Lempel-Ziv-Welch Compression
  - <http://www.geeksforgeeks.org/lzw-lempel-ziv-welch-compression-technique/>
  - <https://en.wikipedia.org/wiki/Lempel-Ziv-Welch>

## 5 Some Suggested Parameters

There will be some discretion in picking implementation parameters. From the start we will suggest considering:

- 4KB average chunk size with a 8KB maximum
- Use on-board DRAM for the chunk dictionary and hash fingerprints
- Full, maximum chunk size as the LZW compression window

We may refine our suggestions as we get more experience, and you may want to experiment with some of the parameters when tuning your implementation.

## 6 Examples of Use

1. Yan Zhang, Nirwan Ansari, Mingquan Wu, and Heather Yu. "On Wide Area Network Optimization." In *IEEE Communications Surveys & Tutorials*, vol. 4, issue 4, pp. 1090–1113, Oct. 2013. <http://ieeexplore.ieee.org/document/6042388/> Sections III A and B survey the role of compression and decompression in optimizing WAN data traffic.
2. Athicha Muthitacharoen, Benjie Chen, and David Mazieres. 2001. A low-bandwidth network file system. In *Proceedings of the eighteenth ACM symposium on Operating systems principles (SOSP '01)*. pp. 174-187. <https://dl.acm.org/citation.cfm?id=502052> Use of deduplication for optimizing a file system operating across a low bandwidth link.

## 7 Chunk Validation

Using an SHA-256 signature, the probability of having a collision where two chunks share the same signature is extremely low. For the project, we will consider equality of SHA-256 signatures adequate to determine that a chunk is a duplicate. This means you do not need to read back the chunk and validate that it is, in fact, identical. If you had terabytes of data, or if the consequences of error were high, you would want to perform the check. This only applies to the full 256b signature. If you use smaller hashes for indexing, you will still need to validate that there is a match on the 256b signature.

## 8 Controlling Heaps and Stacks

The stack and the heap are special memory regions used when processors execute C code. The stack stores values like function parameters and local variables defined within functions. Space on the stack is allocated and freed by compiler automatically. The heap is where you data allocated with `malloc` (and returned with `free`) is stored.

The default stack and heap sizes (0x2000 for stack, 0x10000000 for heap) may be wrong for your project. The default stack size is very small, and you may need more stack space. The default heap is large, and you will probably use very little (if any) dynamically allocated memory. To change the sizes, do the following:

- Go to the SDx *Project Explorer* panel and find the *src* folder under your application. You should find the linker script, `lscript.ld`.
- Double click on the `lscript.ld`, and you will be able to make stack and heap changes here.

## 9 Encoded Data Storage

For your timing runs, store the compressed data in DRAM. The largest test case we provide is under 200MB. Copy your encoded data from DRAM to the SD-Card outside of your test timing. You will need to plan how you divide the DRAM among stack, heap, chunk hash storage, and compressed output. During early testing, before adding ethernet input, you may also want to start with the uncompressed data in DRAM.

## 10 Compressed Format

- Compressed stream is a sequential concatenation of chunks.
- Each chunk has a 32b header that identifies it as Duplicate Chunk or LZW Chunk.
  - A Duplicate Chunk is a 32b value
    - \* bit 0 is a 1 to signify a Duplicate Chunk
    - \* bits 31–1 is the Chunk Index of previously encoded block to be duplicated. Only LZW Chunks are indexed. The first LZW chunk has index 0, the next 1, etc.
  - An LZW Chunk is
    - \* a 32b bit header
      - bit 0 is 0 to signify an LZW Chunk
      - bits 31–1 is the *compressed* chunk length in bytes
    - \* LZW-compressed contents of the chunk. LZW implementations vary. Our implementation satisfies the following properties:
      - Entries 0–255 of the dictionary are initialized to the 256 literals, e.g. a byte with value 27 would be encoded as 27.
      - The next dictionary entries are used for prefixes: sequences of 2 or more bytes.
      - No special keywords such as end-of-file are contained in the dictionary.
      - The dictionary size is only limited by the chunk size limitation.
      - For simplicity, we set all code lengths to be  $\lceil \log_2 \text{MaxChunkSize} \rceil$  in this project. The provided *decoder* works with the  $\text{MaxChunkSize} = 8192$  bytes. You can change the  $\text{MaxChunkSize}$  by changing the defined parameter `CODE_LENGTH`.
      - Code words are output MSB-first. Assuming nothing has been output yet, a 13b code with binary value  $x_{12}x_{11}x_{10}x_9x_8x_7x_6x_5x_4x_3x_2x_1x_0$  results in two consecutive bytes with values  $x_{12}x_{11}x_{10}x_9x_8x_7x_6x_5$  and  $x_4x_3x_2x_1x_0000$ . The next code word with value  $y_{12}y_{11}y_{10}y_9y_8y_7y_6y_5y_4y_3y_2y_1y_0$  changes the second byte to,  $x_4x_3x_2x_1x_0y_{12}y_{11}y_{10}$ , the third to  $y_9y_8y_7y_6y_5y_4y_3y_2$  and the fourth to  $y_1y_0000000$ .
    - \* Padding so that the entire LZW chunk ends on an 8b boundary; that is, chunks of either type always begin on 8b boundaries.

Because we are using a uniform length of  $\lceil \log_2 \text{MaxChunkSize} \rceil > 8$  for encoding codewords, it is possible that an encoded chunk could be longer than the unencoded chunk. To deal with this, real implementations will often compare the length of the LZW encoded chunk to the raw chunk length and send the unencoded data if it is smaller. For simplicity, we are not asking you to perform that optimization (and our encoded chunk format does not support it).

## 11 Ethernet Packet Format and Ethernet MAC Interface

*We will provide information on Ethernet interfacing next week.*

## 12 Supplied Resources

- Design to send data over ethernet at a fixed (tunable) frequency. (*target availability next week*)
- Dummy design to receive ethernet packets from sender. (*target availability next week*)
- [Reference implementation of decoder](#)
- We provide several datasets that you can use for testing. We encourage you to create your own simple datasets for unit testing. Note that the tar-files are not meant to be unpacked. Following are the datasets that we provide.
  - [The Little Prince unencoded](#). As an encoding example, we also provide [The Little Prince compressed](#).
  - [Simple example](#). This archive contains three files, two of which are identical.
  - [Benjamin Franklin's autobiography](#). This is a simple text file that you can modify for your own purposes. The current file probably has few duplicate areas. (390 KB)
  - [GTK+ source code](#). This file contains several subsequent versions of the GTK+ source, which provides ample opportunity for deduplication. (177 MB)
  - [Linux source code](#). This file contains several subsequent versions of the source. (191 MB)
  - [Several Linux kernels](#). As opposed to the other data sets, this set contains prevalently binary data. (66 MB)

Note: you should take these as examples, not a definitive list of test cases. In particular, *you should create many other focused test examples to facilitate your debugging and validation.*

- You can use the compression pipeline of the homeworks as an example of how to access the SD-card.