

Lab 2 Writeup

My name: 李易

My Student number : 191830079

This lab took me about 16 hours to do. I did attend the lab session.

1. Program Structure and Design:

本次实验主要包括两大部分：一个字节流的实现和一个字符串流重组数据结构的实现。

(1) Byte Stream相关

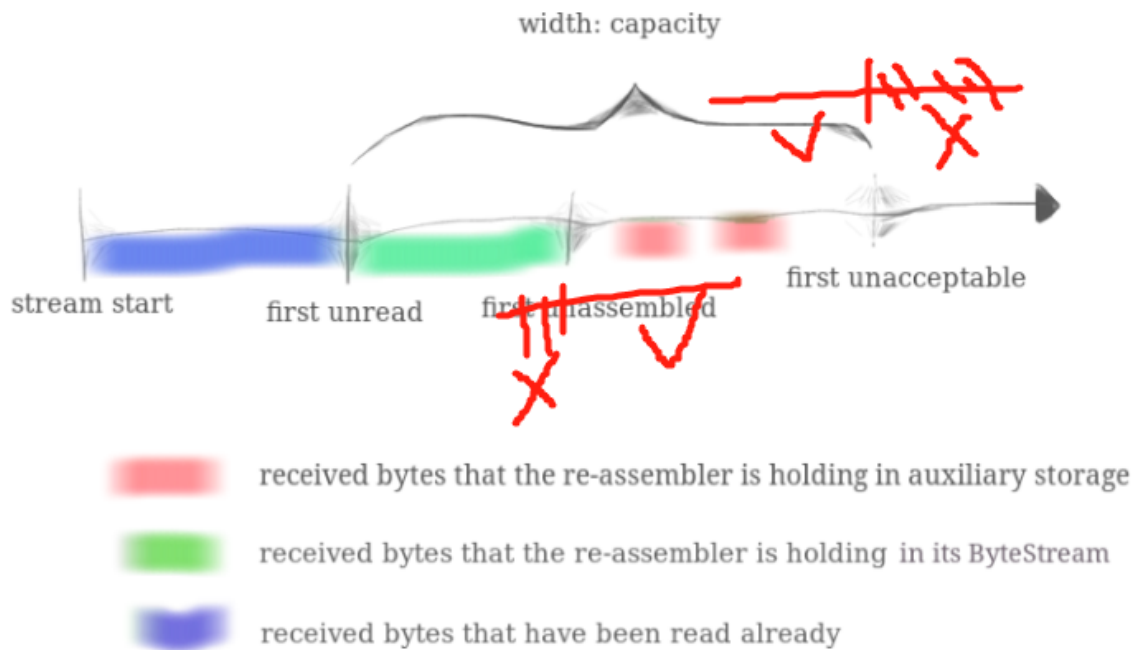
这一部分相对较易，代码量也比较小，主要在于数据结构的选择。

我们的字节流需要有一段写入，一段读取，所以很自然可以想到利用队列去实现。由于之前提到过鼓励大家使用Modern C++代码风格，所以还是尽量用 STL。复习了一下C++相关，由于 `std::queue` 仅支持弹出队首元素，不支持随机访问迭代器等功能，在实现 `peek` 操作时可能比较困难。所以用 `deque` 比较合适。这样的话在 `peek` 操作时可以用迭代器和字符串的 `assign` 函数完成。

(2) Stream Reassembler相关

这部分选择数据结构也很关键。由于我们需要对收到的子串进行重排序、去重合并等功能。所以大概的思路是需要这样一个数据结构，每当我们收到一新子串，按照 `index` 将其插入至正确位置，并利用我们自己写的 `merge_to` 函数将其与其他子串合并，这些子串可能与我们收到的子串交叉。`std::set` 的底层基于红黑树实现，是一种非常高效的关联容器，并且也提供了随机访问迭代器。这里正好可以完成排序、合并的功能。`lower_bound` 接口可以用于快速插入我们新收到的子串。

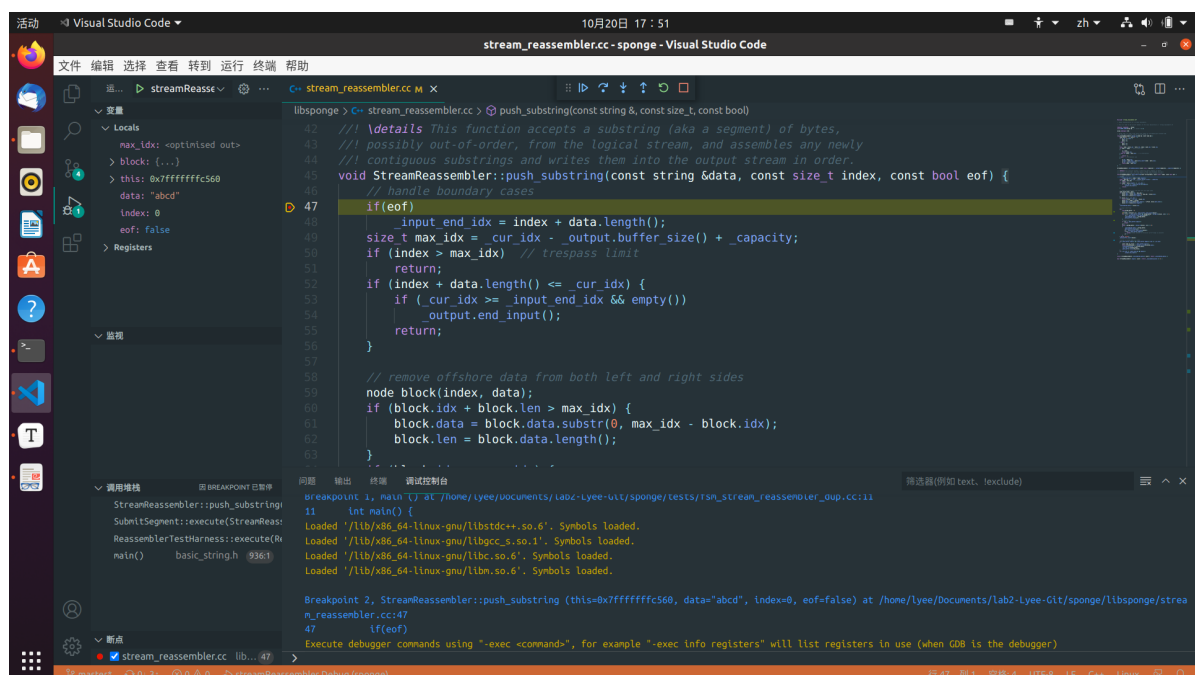
在实现时，我们具体考量几种不同的情况：按序到达时，我们只需讲子串插入 `set`，然后将 `set` 首项立即写入 `byte stream`；若不按序到达，我们将其暂存在 `set` 结构中，等待之后到达的子串与 `set` 中的子串不断合并至 **形成完整的流**后再将其输出到 `byte stream`，并更新相关的变量（当前索引值 `cur_idx`，`_unassembled_bytes` 等）。此时我们依旧只要在对其进行检查（看是否形成了从 `cur_idx` 开始的完整字符串流）后将 `set` 首项输出即可。此外，对于子串中已经写入字节流的部分和超出 `capacity` 的部分应进行截去处理。



所以整个实现大概的流程：

- 1.检查边界情况，看是否超出完全 `capacity` 或者已经 `eof`，这些需要直接返回。
- 2.截去左右两端的越界数据。
- 3.与当前set中已有的 `node` 结构子串合并。
- 4.写入字节流。

2. Implementation Challenges:



利用VSC调试还是很直观的。我主要实现的难点（出现bug的地方）在一些看似比较trivial的地方。也是debug了很久才找到解决方案。

- (1) 对于插入位置之前的 `node` 的相关处理，需要时刻注意迭代器的位置，否则可能越界。
- (2) 进行merge的时候，应该注意好判断时的边界条件，使得正好在形成完整字符串流时两节点能够合并。
- (3) 是否应该eof应在开头和写完 `bytestream` 后各进行一次判断。前者对应输入单个eof的空字符串的情形，后者对应我们写完 `bytestream` 之后恰好完成了所有字符串写入的情形。需要用 `_input_end_index` 记录结束位置，并在每次处理子串时判断是否应该结束输入。

3. Remaining Bugs:

```
lyee@lyee-VirtualBox: ~/Documents/lab2-Lyee-Git/sponge/build
lyee@lyee-VirtualBox:~/Documents/lab2-Lyee-Git/sponge/build$ make check_lab2_1
[100%] Testing Lab 2-part 1...
Test project /home/lyee/Documents/lab2-Lyee-Git/sponge/build
  Start 23: t_byte_stream_construction
1/9 Test #23: t_byte_stream_construction ..... Passed    0.00 sec
  Start 24: t_byte_stream_one_write
2/9 Test #24: t_byte_stream_one_write ..... Passed    0.01 sec
  Start 25: t_byte_stream_two_writes
3/9 Test #25: t_byte_stream_two_writes ..... Passed    0.00 sec
  Start 26: t_byte_stream_capacity
4/9 Test #26: t_byte_stream_capacity ..... Passed    1.20 sec
  Start 27: t_byte_stream_many_writes
5/9 Test #27: t_byte_stream_many_writes ..... Passed    0.01 sec
  Start 28: t_webget
6/9 Test #28: t_webget ..... Passed    0.61 sec
  Start 50: t_address_dt
7/9 Test #50: t_address_dt ..... Passed    0.01 sec
  Start 51: t_parser_dt
8/9 Test #51: t_parser_dt ..... Passed    0.00 sec
  Start 52: t_socket_dt
9/9 Test #52: t_socket_dt ..... Passed    0.01 sec

100% tests passed, 0 tests failed out of 9

Total Test time (real) = 1.87 sec
[100%] Built target check_lab2_1
lyee@lyee-VirtualBox:~/Documents/lab2-Lyee-Git/sponge/build$
```

```
lyee@lyee-VirtualBox: ~/Documents/lab2-Lyee-Git/sponge/build
[100%] Built target address_dt
lyee@lyee-VirtualBox:~/Documents/lab2-Lyee-Git/sponge/build$ make check_lab2_2
[100%] Testing Lab 2-part 2: the stream reassembler...
Test project /home/lyee/Documents/lab2-Lyee-Git/sponge/build
  Start 15: t_strm_reassem_single
1/16 Test #15: t_strm_reassem_single ..... Passed    0.00 sec
  Start 16: t_strm_reassem_seq
2/16 Test #16: t_strm_reassem_seq ..... Passed    0.00 sec
  Start 17: t_strm_reassem_dup
3/16 Test #17: t_strm_reassem_dup ..... Passed    0.02 sec
  Start 18: t_strm_reassem_holes
4/16 Test #18: t_strm_reassem_holes ..... Passed    0.00 sec
  Start 19: t_strm_reassem_many
5/16 Test #19: t_strm_reassem_many ..... Passed    0.18 sec
  Start 20: t_strm_reassem_overlapping
6/16 Test #20: t_strm_reassem_overlapping ..... Passed    0.00 sec
  Start 21: t_strm_reassem_wln
7/16 Test #21: t_strm_reassem_wln ..... Passed    0.17 sec
  Start 22: t_strm_reassem_cap
8/16 Test #22: t_strm_reassem_cap ..... Passed    0.32 sec
  Start 23: t_byte_stream_construction
9/16 Test #23: t_byte_stream_construction ..... Passed    0.00 sec
  Start 24: t_byte_stream_one_write
10/16 Test #24: t_byte_stream_one_write ..... Passed    0.00 sec
  Start 25: t_byte_stream_two_writes
11/16 Test #25: t_byte_stream_two_writes ..... Passed    0.00 sec
  Start 26: t_byte_stream_capacity
12/16 Test #26: t_byte_stream_capacity ..... Passed    1.24 sec
  Start 27: t_byte_stream_many_writes
13/16 Test #27: t_byte_stream_many_writes ..... Passed    0.01 sec
  Start 50: t_address_dt
14/16 Test #50: t_address_dt ..... Passed    0.01 sec
  Start 51: t_parser_dt
15/16 Test #51: t_parser_dt ..... Passed    0.00 sec
  Start 52: t_socket_dt
16/16 Test #52: t_socket_dt ..... Passed    0.01 sec

100% tests passed, 0 tests failed out of 16

Total Test time (real) = 1.99 sec
[100%] Built target check_lab2_2
lyee@lyee-VirtualBox:~/Documents/lab2-Lyee-Git/sponge/build$
```

可能在实现上效率还略有不足？出现了超过0.5 s的测试用例。

