

Java Parallel Streams Internals: Demo'ing Spliterator Performance

Douglas C. Schmidt

d.schmidt@vanderbilt.edu

www.dre.vanderbilt.edu/~schmidt



Professor of Computer Science

**Institute for Software
Integrated Systems**

**Vanderbilt University
Nashville, Tennessee, USA**



Learning Objectives in this Part of the Lesson

- Understand parallel stream internals, e.g.
 - Know what can change & what can't
 - Partition a data source into "chunks"
 - Know the impact of different Java collections on the performance of different spliterators

```
Starting spliterator tests for 10000 words..  
..printing results
```

```
1 msec: LinkedList sequential  
1 msec: ArrayList sequential  
7 msec: ArrayList parallel  
12 msec: LinkedList parallel
```

```
Starting spliterator tests for 100000 words..  
..printing results
```

```
3 msec: ArrayList parallel  
5 msec: ArrayList sequential  
6 msec: LinkedList sequential  
19 msec: LinkedList parallel
```

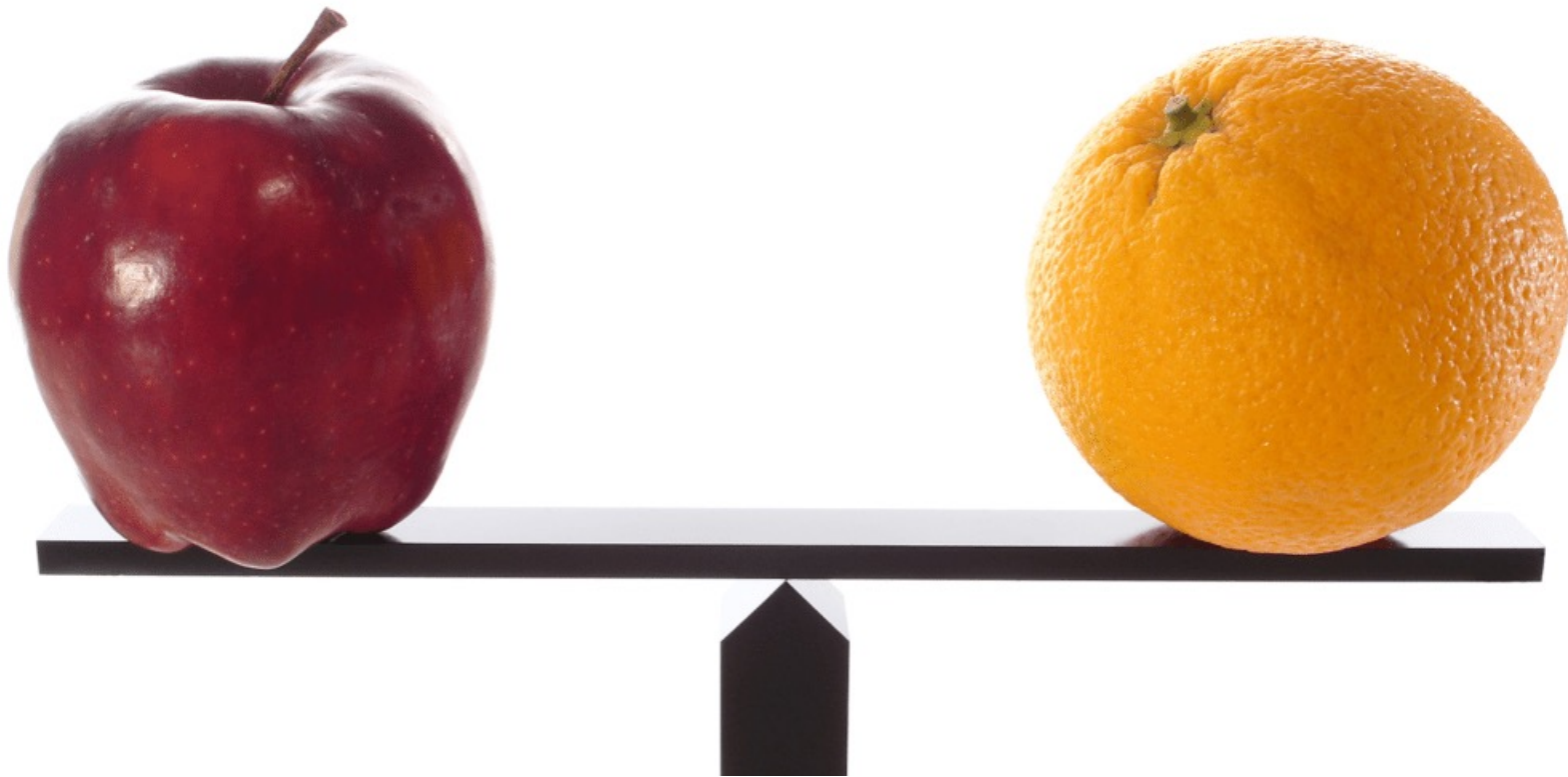
```
Starting spliterator tests for 883311 words..  
..printing results
```

```
12 msec: ArrayList parallel  
14 msec: LinkedList parallel  
38 msec: LinkedList sequential  
43 msec: ArrayList sequential
```

Demonstrating Spliterator Performance

Demonstrating Splitter Performance

- Spliterators for ArrayList & LinkedList partition data quite differently



See earlier lesson on "*Java Parallel Streams Internals: Partitioning*"

Demonstrating Splitter Performance

- Spliterators for ArrayList & LinkedList partition data quite differently

```
ArrayListSpliterator<E> trySplit() {  
    int hi = getFence(), lo = index, mid = (lo + hi) >>> 1;  
    // divide range in half unless too small  
    return lo >= mid ? null : new ArrayListSpliterator<E>  
                                   (list, lo, index = mid, ...);  
}
```

ArrayList's spliterator splits evenly & efficiently (e.g., doesn't copy data)



See openjdk/8u40-b25/java/util/ArrayList.java

Demonstrating Splitterator Performance

- Splitterators for ArrayList & LinkedList partition data quite differently

```
Splitterator<E> trySplit() { ...  
    int n = batch + BATCH_UNIT, j = 0; Object[] a = new Object[n];  
    do { a[j++] = p.item; }  
    while ((p = p.next) != null && j < n); ...  
    return Spliterators.splititerator(a, 0, j, Splitterator.ORDERED);  
}
```

LinkedList's splitterator does not split evenly & efficiently (e.g., it copies data)



See openjdk/8u40-b25/java/util/LinkedList.java

Demonstrating Splitter Performance

- This demo program shows the performance difference of parallel spliterators for ArrayList & LinkedList when processing the complete works of Shakespeare

```
void timeStreamUppercase
    (String testName, List<String> words, boolean parallel) {
    ...
    List<String> list = new ArrayList<>();

    IntStream.range(0, sMAX_ITERATIONS).forEach((i) -> list
        .addAll(StreamSupport
            .stream(words.splitIterator(),
                parallel)
            .map(string -> string.toUpperCase())
            .toList()));

    ...
}
```

Demonstrating Splitter Performance

- This demo program shows the performance difference of parallel spliterators for ArrayList & LinkedList when processing the complete works of Shakespeare

```
void timeStreamUppercase
(String testName, List<String> words, boolean parallel) {
    ...
    List<String> list = new ArrayList<>();
    IntStream.range(0, sMAX_ITERATIONS).forEach((i) -> list
        .addAll(StreamSupport
            .stream(words.splititerator(),
                parallel)
            .map(string -> string.toUpperCase())
            .toList()));
    ...
}
```

The words param is passed an ArrayList & a LinkedList

Demonstrating Splitter Performance

- This demo program shows the performance difference of parallel spliterators for ArrayList & LinkedList when processing the complete works of Shakespeare

```
void timeStreamUppercase
(String testName, List<String> words, boolean parallel) {
    ...
    List<String> list = new ArrayList<>();

    IntStream.range(0, sMAX_ITERATIONS).forEach((i) -> list
        .addAll(StreamSupport
            .stream(words.splitter(),
                parallel)
            .map(string -> string.toUpperCase())
            .toList()));
    ...
}
```

*Split & uppercase the
word list via a splitter*

Demonstrating Spliter Performance

- This demo program shows the performance difference of parallel spliterators for ArrayList & LinkedList when processing the complete works of Shakespeare

```
void timeStreamUppercase
(String testName, List<String> words, boolean parallel) {
    ...
    List<String> list = new ArrayList<>();

    IntStream.range(0, sMAX_ITERATIONS).forEach((i) -> list
        .addAll(StreamSupport
            .stream(words.splitter(),
                parallel)
            .map(string -> string.toUpperCase())
            .toList()));
    ...
}
```

*Conditionally select a parallel
or sequential splitter*

Demonstrating Splitter Performance

- Results show splitter differences become more significant as input grows

...

Starting splitter tests for 100000 words..

..printing results

3 msecs: ArrayList parallel

5 msecs: ArrayList sequential

6 msecs: LinkedList sequential

19 msecs: LinkedList parallel

Starting splitter tests for 883311 words..

..printing results

12 msecs: ArrayList parallel

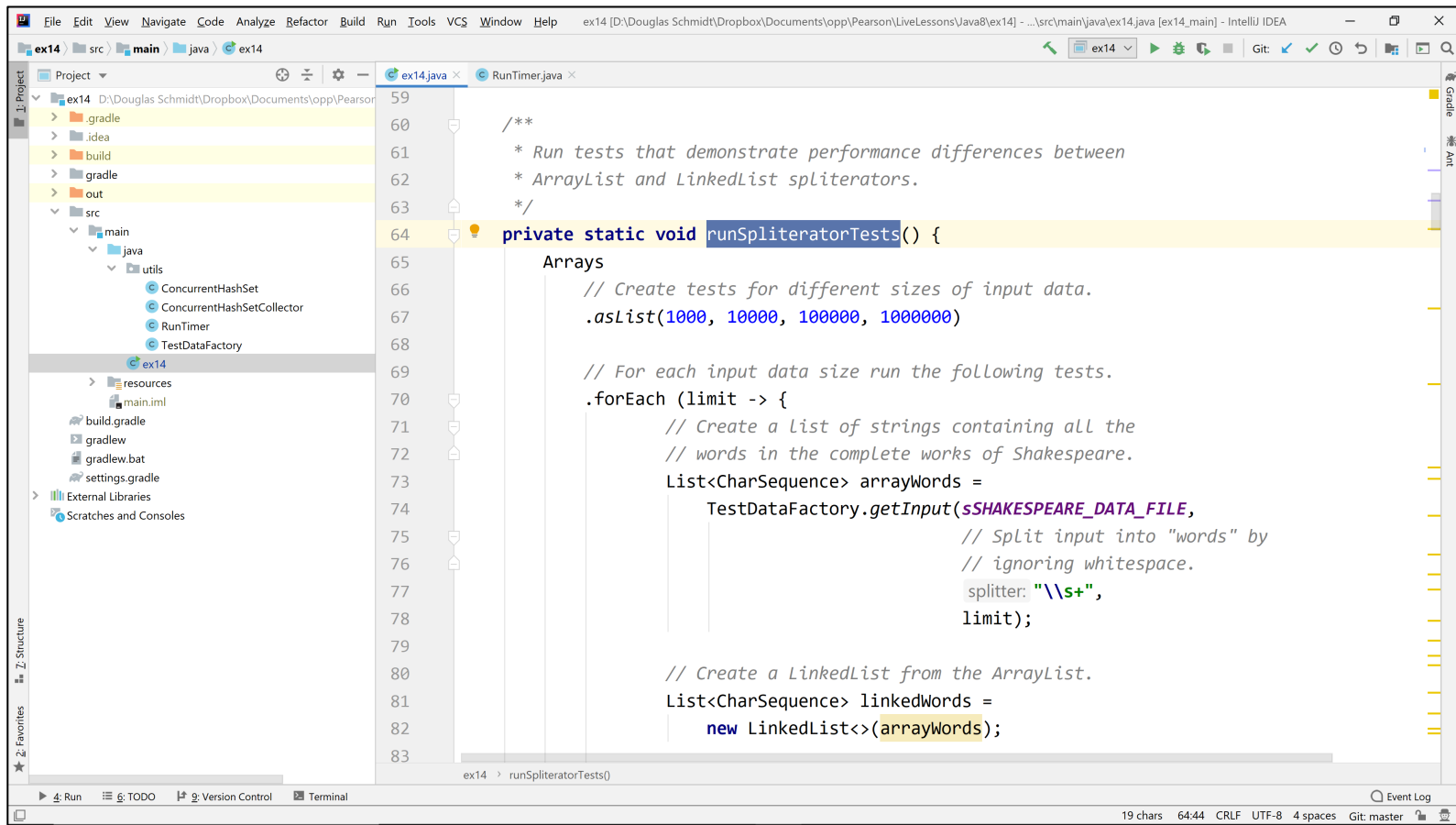
14 msecs: LinkedList parallel

38 msecs: LinkedList sequential

43 msecs: ArrayList sequential

See upcoming lessons on "*When [Not] to Use Parallel Streams*"

Demonstrating Spliterator Performance



```
59
60 /**
61  * Run tests that demonstrate performance differences between
62  * ArrayList and LinkedList spliterators.
63  */
64 private static void runSpliteratorTests() {
65     Arrays
66     // Create tests for different sizes of input data.
67     .asList(1000, 10000, 100000, 1000000)
68
69     // For each input data size run the following tests.
70     .forEach (limit -> {
71         // Create a list of strings containing all the
72         // words in the complete works of Shakespeare.
73         List<CharSequence> arrayWords =
74             TestDataFactory.getInput(SSHAKESPEARE_DATA_FILE,
75                                     // Split input into "words" by
76                                     // ignoring whitespace.
77                                     splitter: "\\s+",
78                                     limit);
79
80         // Create a LinkedList from the ArrayList.
81         List<CharSequence> linkedWords =
82             new LinkedList<>(arrayWords);
83     })
```

See github.com/douglasraigschmidt/LiveLessons/tree/master/Java8/ex14

End of Java Parallel Streams Internals: Demo'ing Spliterator Performance