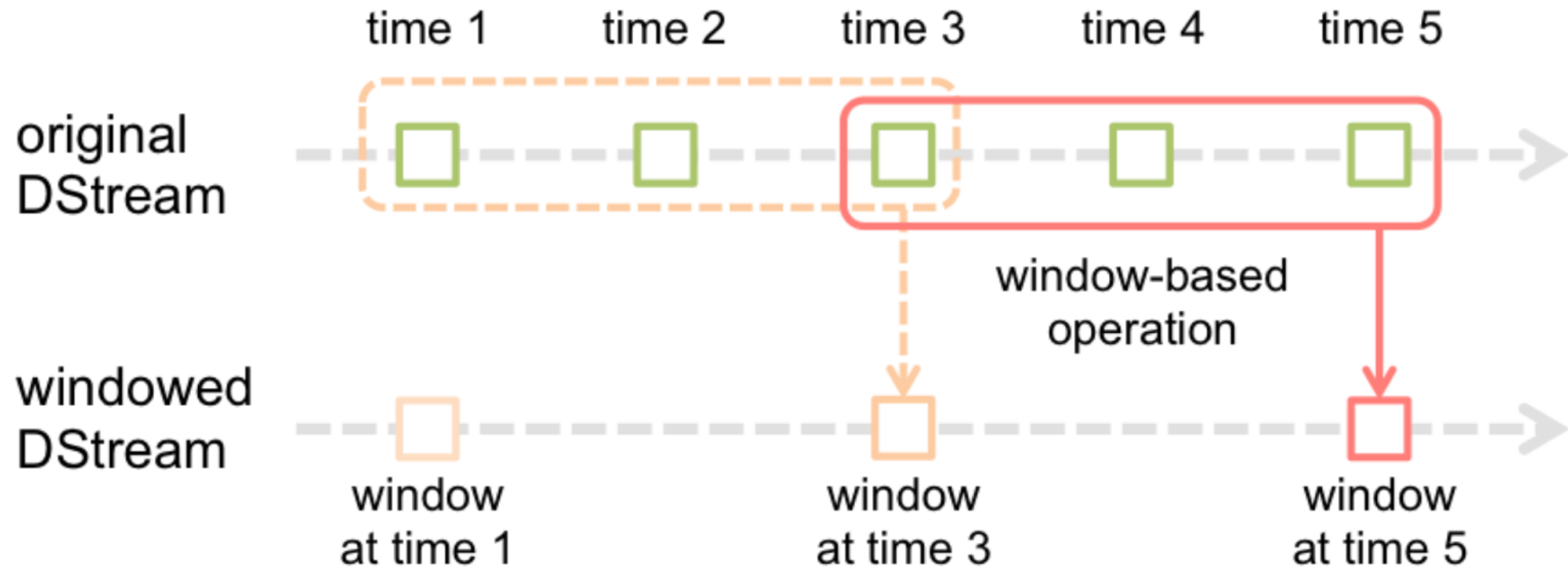


# Spark Streaming Windows

# Windows



- window length - The duration of the window (3 in the figure).
- sliding interval - The interval at which the window operation is performed (2 in the figure).

These two parameters must be multiples of the batch interval of the source DStream (1 in the figure).

# Transformation with Window

```
words = lines.flatMap(lambda line: line.split(" "))
pairs = words.map(lambda word: (word, 1))
pairs.window(30, 10).pprint()
```

```
ssc.start()
```

```
-----  
Time: 2018-12-31 00:28:55  
-----
```

```
('bbb', 1)  
( 'bbb', 1)  
( 'bbb', 1)  
( 'cccb', 1)  
( 'c', 1)  
( 'cc', 1)  
( 'cccc', 1)  
  
-----
```