Output Operations on DStreams

Overview

- Output operations allow DStream’s data to be pushed out to external systems like a database or a file systems.
- Since the output operations actually allow the transformed data to be consumed by external systems, they trigger the actual execution of all the DStream transformations (similar to actions for RDDs).

Currently Defined Output Operations

Output Operation	Meaning
`pprint()`	Prints the first ten elements of every batch of data in a DStream on the driver node running the streaming application. This is useful for development and debugging.
`saveAsTextFiles(`*`prefix`*`, [`*`suffix`*`])`	Save this DStream's contents as text files. The file name at each batch interval is generated based on prefix and suffix: "prefix-TIME_IN_MS[.suffix]".
`saveAsObjectFiles(`*`prefix`*`, [`*`suffix`*`])`	Save this DStream's contents as SequenceFiles of serialized Java objects. The file name at each batch interval is generated based on prefix and suffix: "prefix-TIME_IN_MS[.suffix]".
`saveAsHadoopFiles(`*`prefix`*`, [`*`suffix`*`])`	Save this DStream's contents as Hadoop files. The file name at each batch interval is generated based on prefix and suffix: "prefix-TIME_IN_MS[.suffix]".
`foreachRDD(`*`func`*`)`	The most generic output operator that applies a function, func, to each RDD generated from the stream. This function should push the data in each RDD to an external system, such as saving the RDD to files, or writing it over the network to a database.
* Greyed out indicates no support from Python API	

To the Code!