To be honest I would like to end my answer by saying that Apache Storm is an excellent state of the art stream processing framework which allows us to do *extreme low level latency event* processing (Core Storm) and micro-batch processing (Trident Storm) *for higher throughput*.

<http://storm.apache.org/releases/1.0.1/Trident-tutorial.html>

Ok. **Spark** is a distributed batch processing system that supports micro-batching-based stream processing. Batches are distributed across a cluster and reduced down to a result.

**Storm** is a distributed stream processing system that supports micro-batching. Streams are routed through bolts and spouts which you assemble yourself manually or using a higher level, extraordinarily badly documented framework called trident.

**Flink** is a stream processing engine on top of which micro-batch and batch processing engines have been built. Like spark, it supports iterative computation, which makes it potentially suitable for a broad range of uses cases.

**Beam** is a library for specifying data flows that can be executed on any of the above framework. E.g. you can write a beam flow on spark or on flink or potentially on storm.

**Apex** is the core of the DataTorrent RTS system and also supports streaming and micro-batching. I'm not personally familiar with its architecture so that's all I can really say.

Conclusion: The big data engine ecosystem is extremely active these days! Also, quite confusing.