**CHAPTER 1**

**Map Reduce**

|  |
| --- |
| Storm is a distributed, reliable, fault-tolerant system for processing streams of data.  The work is delegated to different types of components that are each responsible for a simple specific processing task. The input stream of a Storm cluster is handled by a component called a spout. The spout passes the data to a component called a bolt, which transforms it in some way. A bolt either persists the data in some sort of storage, or passes it to some other bolt. |

*“I wondered if each name was mentioned an equal number of times, or if there was a bias in the number of mentions.”*

Use Cases:

|  |  |  |
| --- | --- | --- |
| Processing stream | Continuous computation | Distributed remote procedure call |