# HBASE WRITE AHEAD LOG PERFORMANCE EVALUATION

*A Performance evaluation for Open Source BigTable Implementations. Ankur Khetrapal . split the HRegionServer's write-ahead log so that there is now one.*

Let"s look at the high level view of how this is done in HBase. HBase is very much a developer-centric database. By default importtsv will load data directly into HBase. An HBase master node serves a Web interface on port  But Hadoop is really an umbrella name for an entire ecosystem of technologies, some of which HBase uses to create a distributed, column-oriented database built on the same principles as Google's Bigtable. The HBase endpoint coprocessor works much like a stored procedure. The used SequenceFile has quite a few shortcomings that need to be addressed. That is also why the downward arrow in the big picture above is done with a dashed line to indicate the optional step. HBase ships another diagnostic mapreduce job called CellCounter. Avro is also slated to be the new RPC format for Hadoop, which does help as more people are familiar with it. However, HDFS provides mechanisms that advertise block location and -- more important -- perform block relocation upon request. We are talking about fsync style issues. Multiple nodes can and should be designated as master nodes, but when the cluster boots, the candidate masters coordinate so that only one is the acting master. If you want to understand the role played by a particular HBase entity -- say, a Filter -- be prepared to be handed off to the Java API's documentation of the Filter class for a full explanation. As a result, all random writes are performed in memory, and when data is flushed to disk, the data is first sorted, then written sequentially with an accompanying index. You can increment any number of counters within a single row via a single call, and without having to lock the row. You will see in a minute where this is used. Other options that may be specified with -D include: -Dimporttsv. The output can optionally be mapped to another set of tables. Use hbase.