Hadoop Is Falling – Why?

Three years ago, looking beyond Hadoop was insanity, and there was little else that could come close. Recently, adoption of Hadoop has slowed down considerably. We examine why.

By George Hill, Innovation Enterprise.

Adoption has slowed considerably, but is that the fault of the platform?

Three years ago, looking beyond Hadoop was insanity, and there was little else that could come close according to many in the media. However, the reality has been a little different.

For a long period, Hadoop and big data were almost interchangeable when they were being discussed by those in the media, although this was not necessarily found to be the case amongst data scientists. A study by Silicon Angle in 2012 analyzing Twitter conversations between data professionals talking about big data found that they actually talked about NoSQL technologies like MongoDB as much, or more, than Hadoop, which would indicate that it has not actually been the must have that many assumed it was.

Most would argue that Hadoop has been one of the single most important elements in the spread of big data, that it is very much the foundation on which data today is built. We are also still finding new ways to use it, in warehousing for instance. That being

said, to the surprise of many, its adoption appears to have more or less stagnated, leading even James Kobielus, Big Data Evangelist at IBM Software, to claim that 'Hadoop declined more rapidly in 2016 from the big-data landscape than I expected.'

The reasons for this are hard to ascertain, but could be down to a problem common in data circles. A 2015 study from Gartner found that 54% of companies had no plans to invest in Hadoop, while 44% of those asked had adopted Hadoop already or planned to at some point in the next two years. This could, depending on your point of view, be taken to mean either that it would see even further expansion or that the majority were ignoring it. However, the survey also revealed a number of other telling factors with implications unlikely to have subsided since. Of those who were not investing, 49% were still trying to figure out how to use it for value, while 57% said that the skills gap was the major reason, a number that is not going to be corrected overnight. This coincides with findings from Indeed who tracked job trends with 'Hadoop Testing' in the title, with the term featured in a peak of 0.061% of ads in mid 2014, which then jumped to 0.087% in late 2016, an increase of around 43% in 18 months.

What this may signal is that adoption hasn't necessarily dropped to the extent that anecdotal evidence would suggest, but companies are simply finding it difficult to extract value from Hadoop from their current teams and they require greater expertise.

Another element that may be cause for concern is simply that one man's big data is another man's small data. Hadoop is designed for huge amounts of data, and as Kashif Saiyed wrote on KD Nuggets 'You don't need Hadoop if you don't really have a problem of huge data volumes in your enterprise, so hundreds of enterprises were hugely disappointed by their useless 2 to 10TB Hadoop clusters – Hadoop technology just doesn't shine at this scale.'

Most companies do not currently have enough data to warrant a Hadoop rollout, but did so anyway because they felt they needed to keep up with the Joneses. After a few years of experimentation and working alongside genuine data scientists, they soon realize that their data works better in other technologies.

This trend has had impacts beyond a slow down in the adoption of an open source platform though, for some companies this has had real world financial impacts. Cloudera and Hortonworks are two of the biggest companies that build their products out from a Hadoop framework. Both have lost significant value in-part due to its decline, with Cloudera reported to have lost 40% whilst Hortonworks' shares have plummeted 68% since mid 2015.

Criticism within this article may seem harsh on Hadoop, but it is not the platform in itself that has caused the current issues. Instead it is perhaps the hype and association

of big data that has done the real damage. Companies have adopted the platform without understanding it and then failed to get the right people or data to make it work properly, which has led to disillusionment and its apparent stagnation. There is still a huge amount of life in Hadoop, but people just need to understand it better.

Original post. Reposted with permission.

Bio: George Hill is the Editor-In-Chief at Innovation Enterprise, Co Founder at The Cyclist.

Related:

- Enterprise Hadoop Adoption: Half Empty or Half Full?, from Hortonworks.
- Why Not So Hadoop?
- Hadoop and Big Data: The Top 6 Questions Answered
- Hadoop Key Terms, Explained