

# ***How Do You Hadoop?***

*An update to “Do You Hadoop?”*

**May 30, 2014**

**Shirish Netke  
M. R. Rangaswami**

**SandHill**  
Group

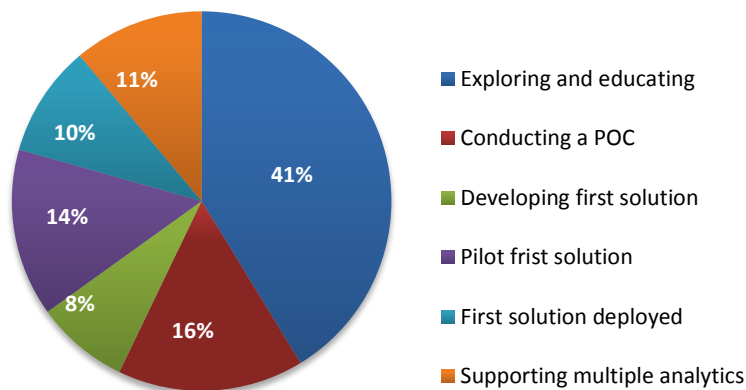
## Executive Summary – *How Do You Hadoop?*

There has been a lot of activity around Hadoop offerings in the last six months. This is quick snapshot of how Hadoop practitioners' responses have changed since our last report "Do You Hadoop?" in 2013. It is based on a survey of 63 companies in May 2014. Here are the highlights:

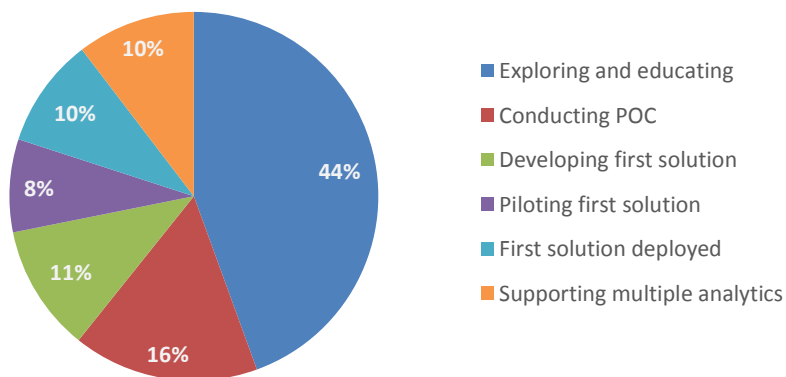
- 1. More Hadoop pilots are underway**
- 2. Increased variety and complexity of data**
- 3. Satisfaction with Hadoop initiatives remains high**
- 4. Skills shortage in MapReduce / Statistics / SQL-on-Hadoop**
- 5. Building momentum around Hadoop 2.0/YARN**

# More Hadoop Pilots Underway

May 2014



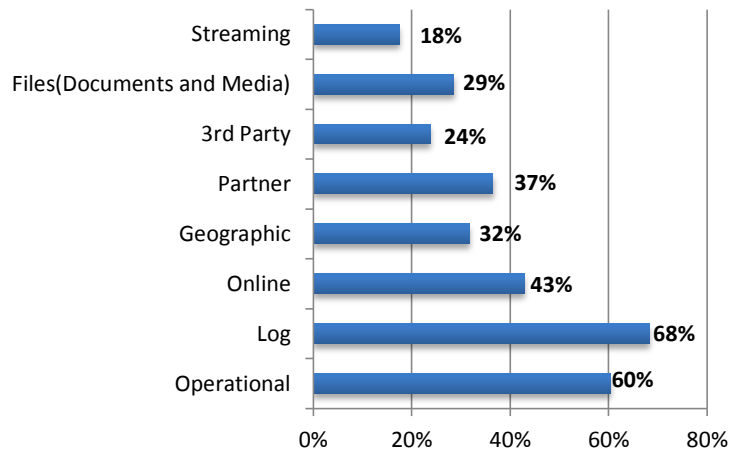
October 2013



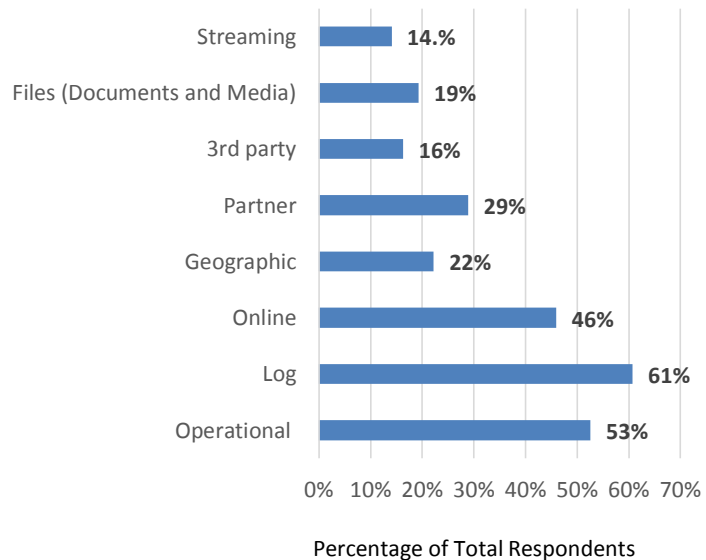
- Pilot activity on Hadoop projects has increased significantly with 14% reporting their first pilot solution up from 8% in October.
- Slight reduction in the percentages of early-stage activities such as exploration and first solution development.
- Production deployments similar to before but may rise as pilots are completed.

# Increased Variety and Complexity of Data

**May 2014**



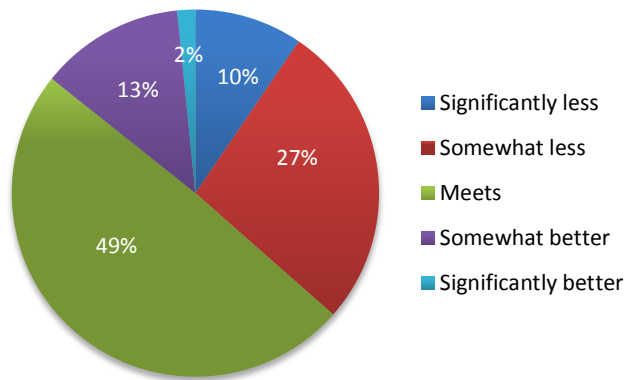
**October 2013**



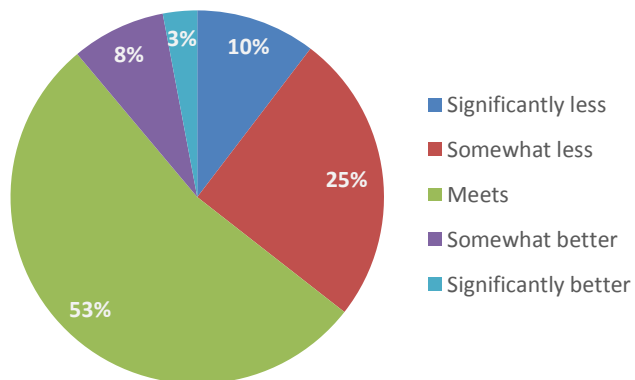
- Significant increase in the usage of Hadoop across the board for most data types since October.
- The substantial increase in third-party data and geographical data implies an increased confidence in handling complex data types that require advanced technical skills.

# Satisfaction with Hadoop Initiatives Remains High

## Overall Satisfaction May 2014



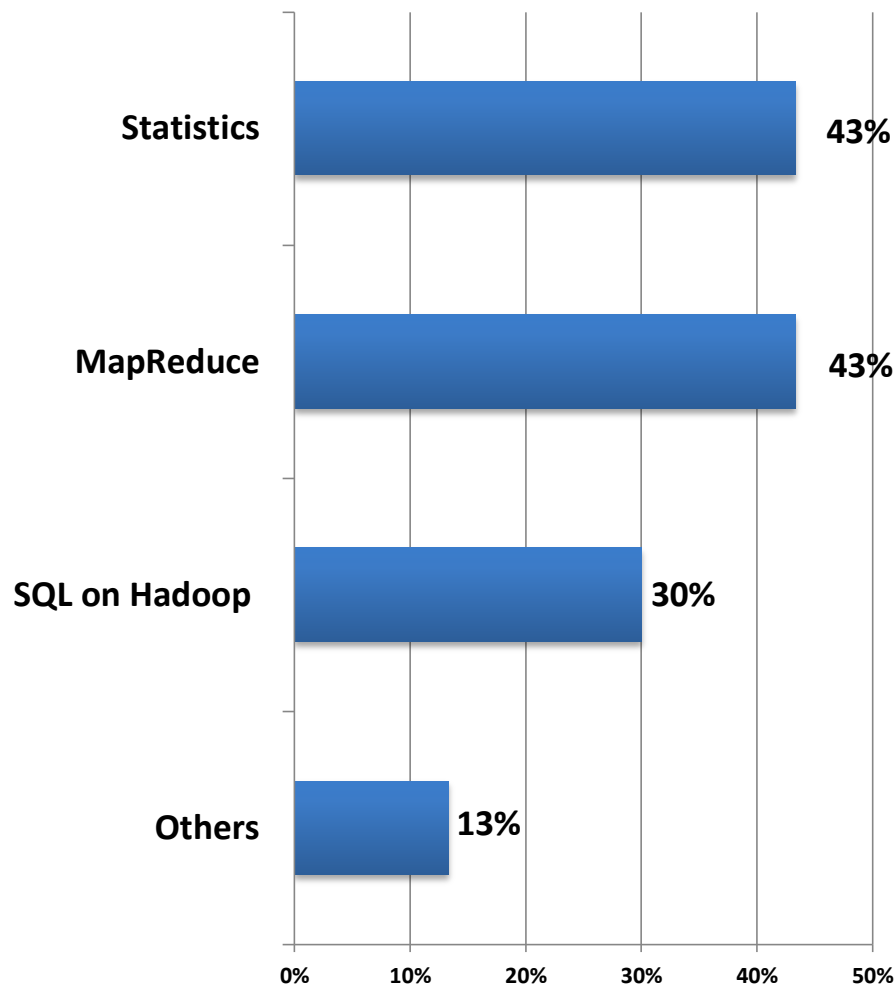
## Overall Satisfaction October 2013



- Hadoop has met or exceeded the expectations of 64% of the users in May 2014, which is almost the same as it was in October.
- A high level of satisfaction on Hadoop initiatives is consistent with the increased focus on more complex and advanced Hadoop initiatives.

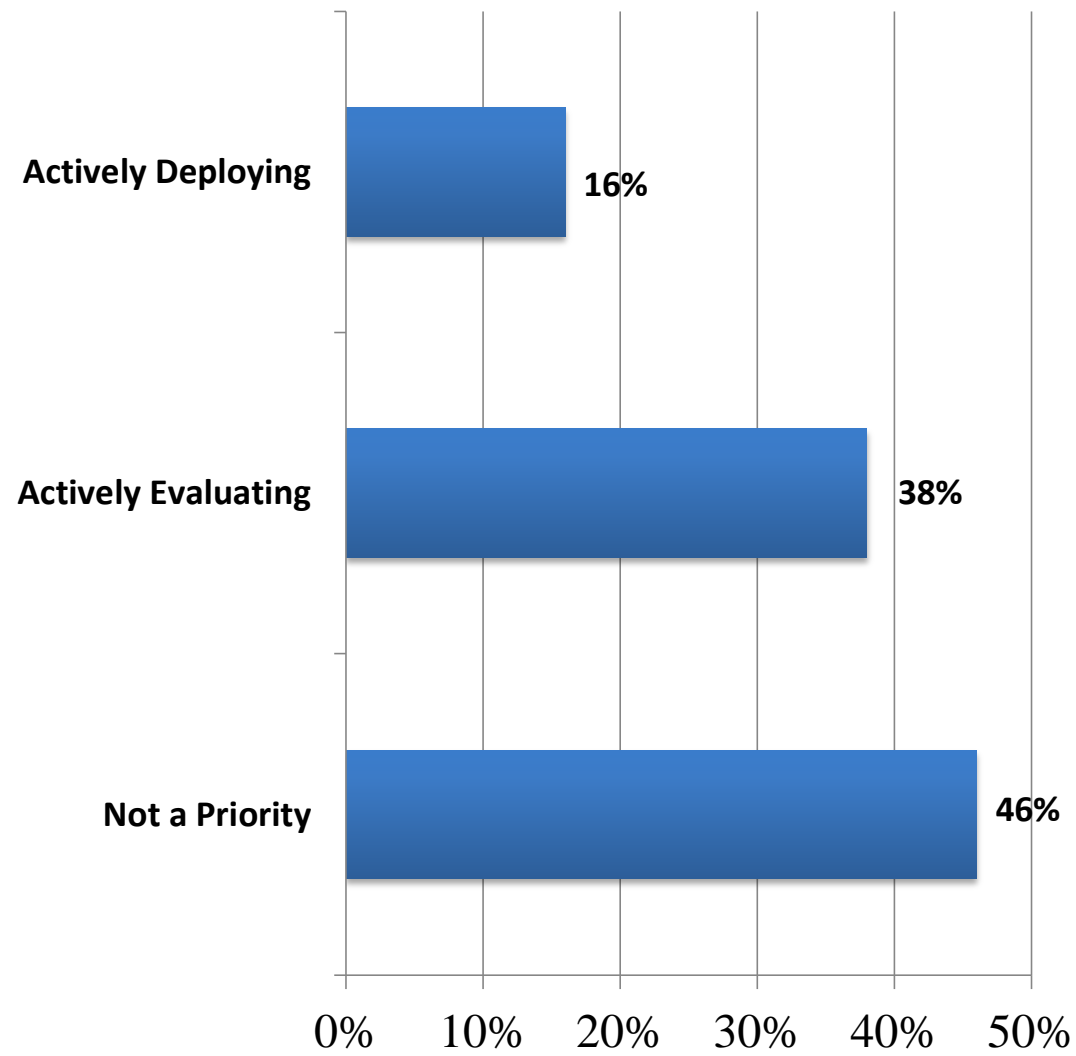
# Skills Shortage in Statistics / MapReduce / SQL-on-Hadoop

Shortage Skills May 2014



- Shortage of statistics skills is as critical as the shortage of MapReduce skills
- SQL-on-Hadoop gaining importance
  - 22% had already deployed a SQL-on-Hadoop solution
  - 40% are evaluating a SQL-on-Hadoop solution
  - 30% cited SQL-on-Hadoop as a shortage skill.
- Configuration and support highlighted as additional shortage skills

# Building Momentum around Hadoop 2.0/YARN



- A majority of the respondents (53%) are evaluating or have deployed Hadoop 2.0/YARN
- 64% of respondents who were at or beyond the POC stage are evaluating or have deployed Hadoop 2.0/YARN
- More than 35% of respondents who are in the exploration and education phase are evaluating Hadoop 2.0/YARN

# Shirish Netke

*Shirish Netke has extensive leadership experience in high-growth technology companies. He has been quoted as a thought leader in the New York Times, Investors Business Daily, Chief Executive Magazine and Asia Times. His speaking engagements have covered a variety of forums such as Seoul Digital Forum, GITEX, CFO Rising and Oracle OpenWorld. He is currently the CEO and President of Amberoon, a provider of data-driven business perspective solutions. Contact [shirish.netke@amberoon.com](mailto:shirish.netke@amberoon.com).*



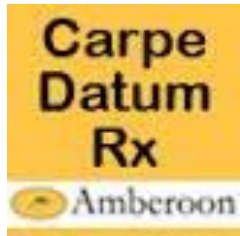


# M. R. Rangaswami

*M.R. Rangaswami thrives on creating influential communities that achieve global impact. Using his skills as a critic, cheerleader and facilitator, M.R. has united leaders and advanced strategic initiatives in the fields of software, corporate sustainability and entrepreneurship. He is the publisher of SandHill.com, the premier online destination for strategic information on the Big Data, cloud and mobile software ecosystem. Contact M.R. at [mr@sandhill.com](mailto:mr@sandhill.com).*



# Partners



*Carpe Datum Rx is a thought leadership forum for the business application of advanced analytics technologies.*

[www.CarpeDatumRx.com](http://www.CarpeDatumRx.com)



*The Hive works with entrepreneurs to create companies that use data to innovate products and gain advantage in a competitive market.*

[www.HiveData.com](http://www.HiveData.com)