

451

Research

ANALYZING THE BUSINESS
OF ENTERPRISE IT INNOVATION

Avoid Re-Inventing the Wheel When Seeking Big Data Bliss

April 9th, 2014

NEW YORK · BOSTON · WASHINGTON DC · SAN FRANCISCO · SEATTLE · DENVER · LONDON · SAO PAULO · DUBAI · SINGAPORE

WWW.451RESEARCH.COM



Michael Coté
Research Director,
Infrastructure

ch.com

@cote - <http://cote.io>

Responsible for systems management, application development, cloud software, and misc. “infrastructure software” agenda

Worked at Dell in corporate strategy, as an analyst for 6+ years, software developer for 10+ years



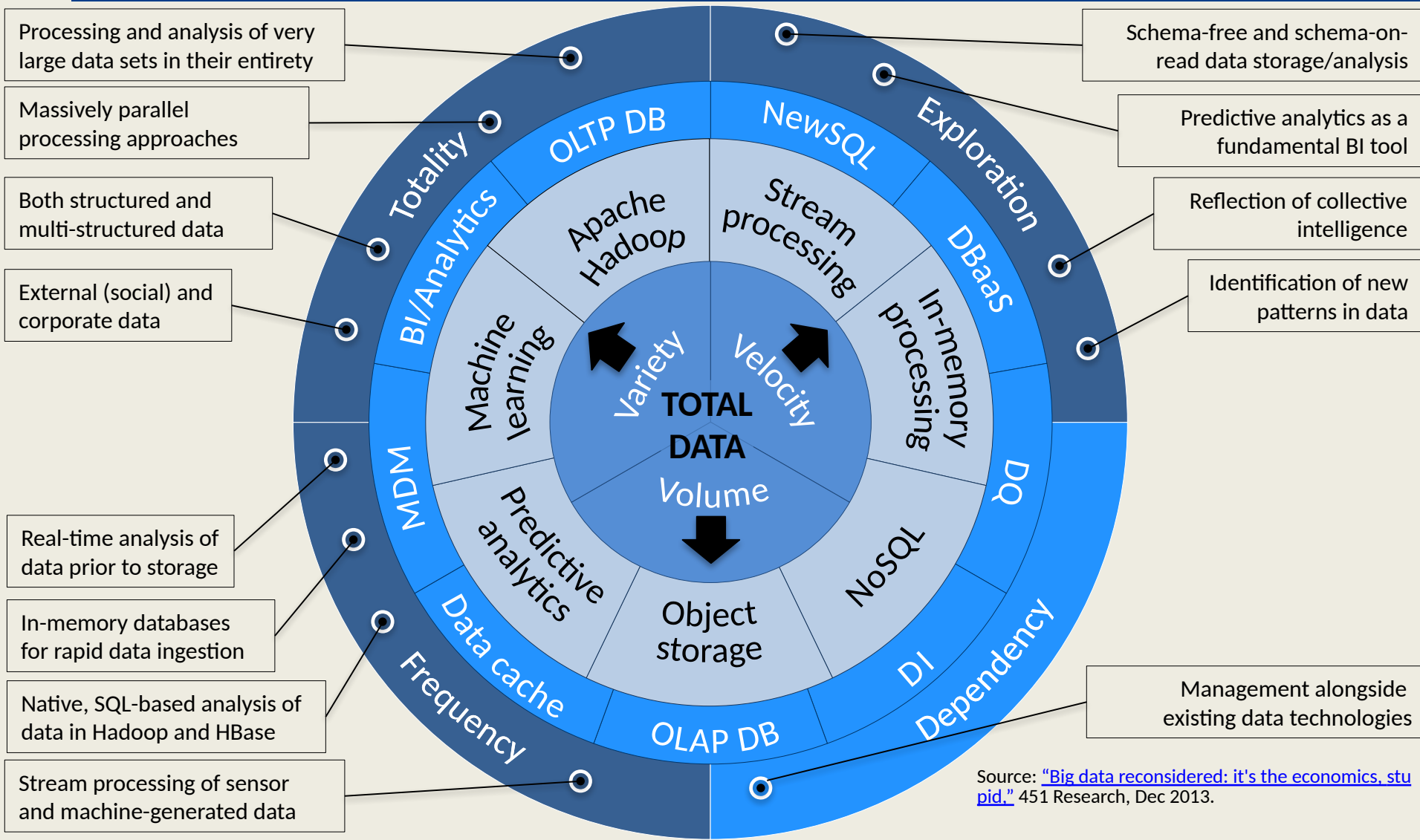
Joe Goldberg
BMC Control-M
Solutions Marketing
joe_goldberg@bmc.com
[@GoldbergJoe](#)

Joe is an IT professional with over 35 years of experience in the design, development, implementation, sales and marketing of enterprise solutions to Global 2000 organizations. Joe has been active in helping BMC products leverage new technology to deliver market-leading solutions

BMC slides were omitted from this presentation. See full presentation and recording here:

<https://www.brighttalk.com/webcast/9059/103135>

Usually, "Big Data" us a synonym for "Hadoop:" not so fast



Source: "[Big data reconsidered: it's the economics, stupid.](#)" 451 Research, Dec 2013.

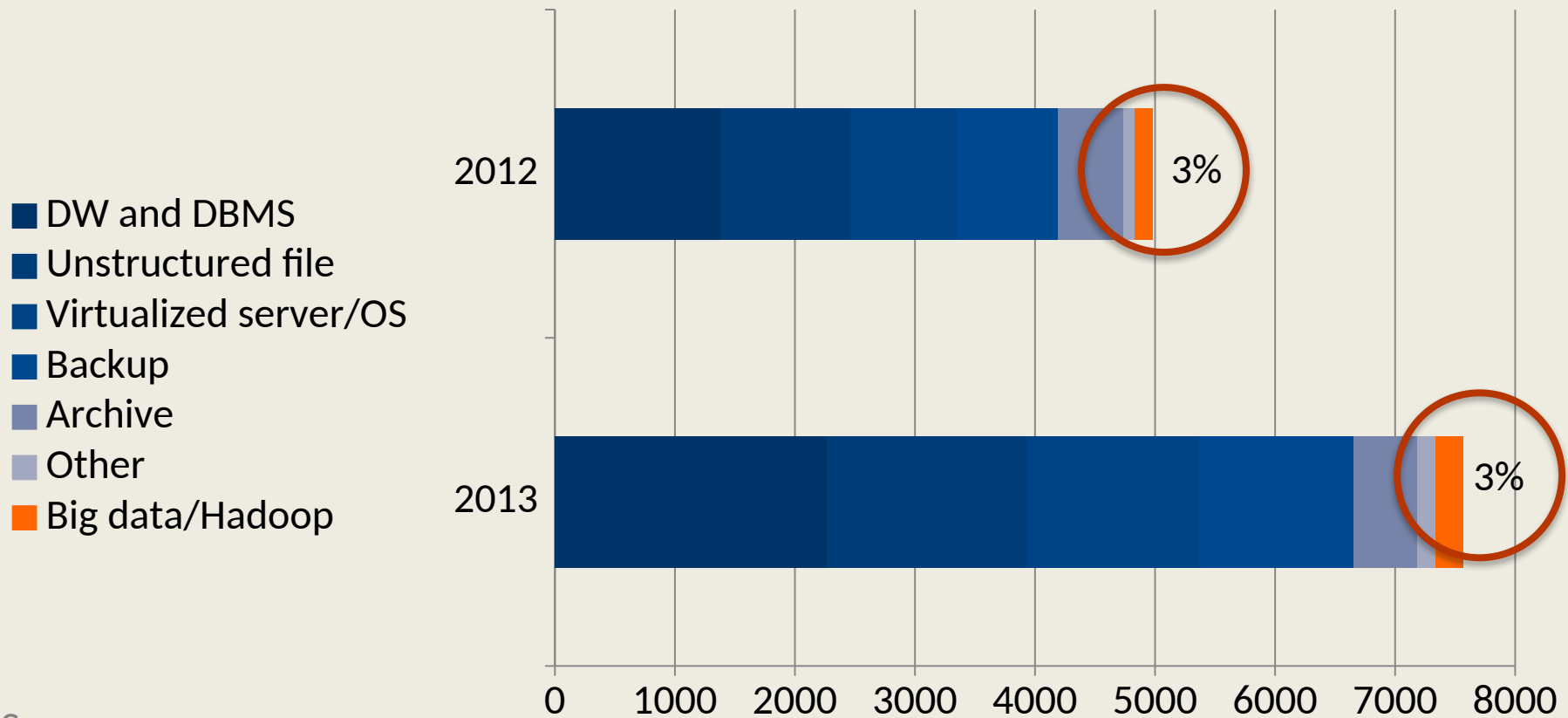
Big Data: “it’s the economics stupid”

- **“The price point that Hadoop comes in at is transformational. Hadoop has the ability to drive down operational cost and improve resource efficiency.”**
 - **Global Head of Architecture, Global Bank**

- Another example: a provider of real-time information and analysis to the media and communications industries
 - Moved from storing 1% of data for 60 days in EDW @ \$100,000/TB
 - To 100% of data for a year in Hadoop @ \$900/TB
 - By migrating to Hadoop and open source databases the company identified over \$4m in cost savings over two years
- Both companies have retained the use of traditional databases/warehousing, but Hadoop and other big data technologies add cost-effectiveness and flexibility

'Big data' not significant in core infrastructure yet

Average total storage capacity (TBs), and total storage footprint by workload illustrate the low level of adoption at today



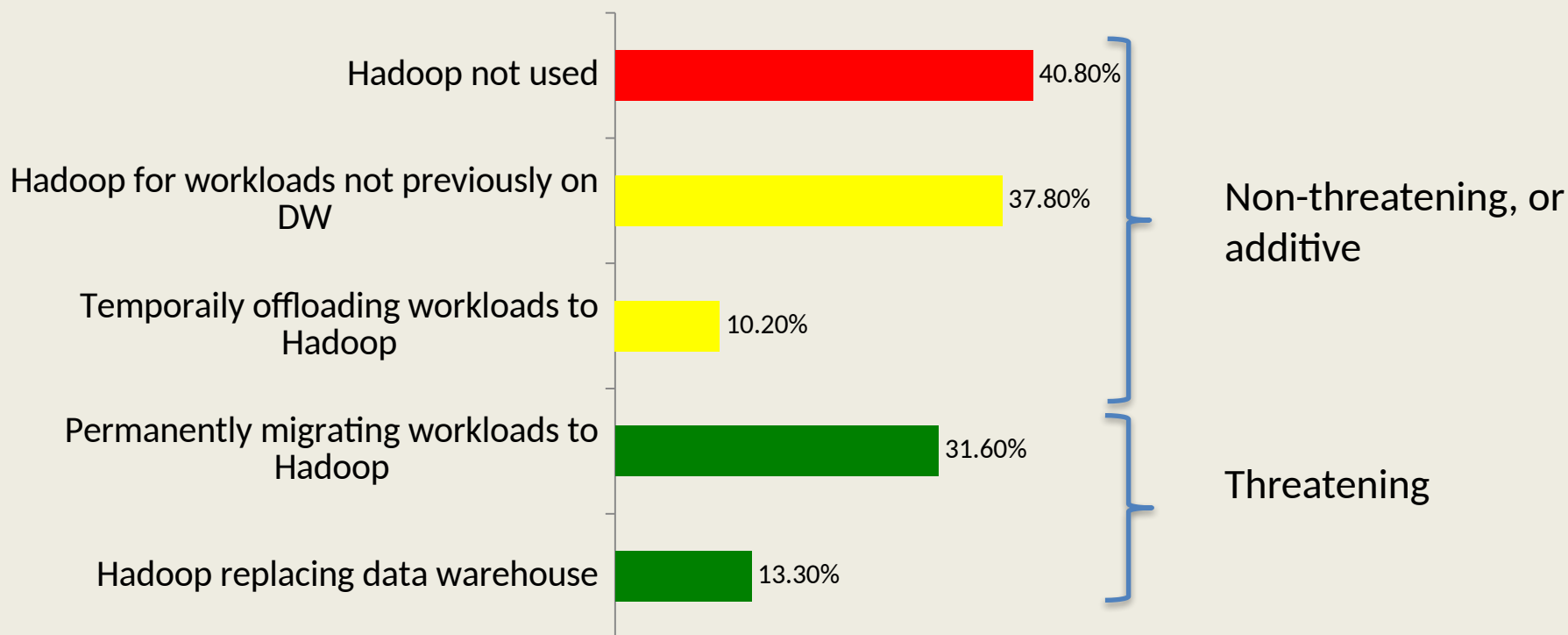
Source:

2012: 451 Research The Info Pro Storage - Wave 16 | n=214

2013: 451 Research The Info Pro Storage - Wave 17 | n=200

Hadoop vs. EDW – not so much

Describe the relationship between Hadoop and the data warehouse within your organization



Source: ["Hadoop: a framework in search of a metaphor."](#) 451 Survey conducted Sep/Oct 2013, sample=98.

What's big data good for?

- The processing and analysis of very large data sets in their entirety
- Increased adoption of massively parallel processing approaches
- Storage and analysis of both structured and un-structured data
- Integration of external (social) and corporate data for more complete perspective
- *Ad hoc* analytic approaches to identify new patterns in data
- Interactive, native, SQL-based analysis of data in Hadoop and Hbase.
- Predictive analytics as a fundamental component of BI strategies
- Machine-learning algorithms automate the reflection of collective intelligence
- Increased adoption of in-memory databases for rapid data ingestion
- Stream processing of sensor and other machine-generated data/events
- Real-time analysis of data prior to storage within the data warehouse/Hadoop
- “MR-ETL” – pre-processing data for EDW loads

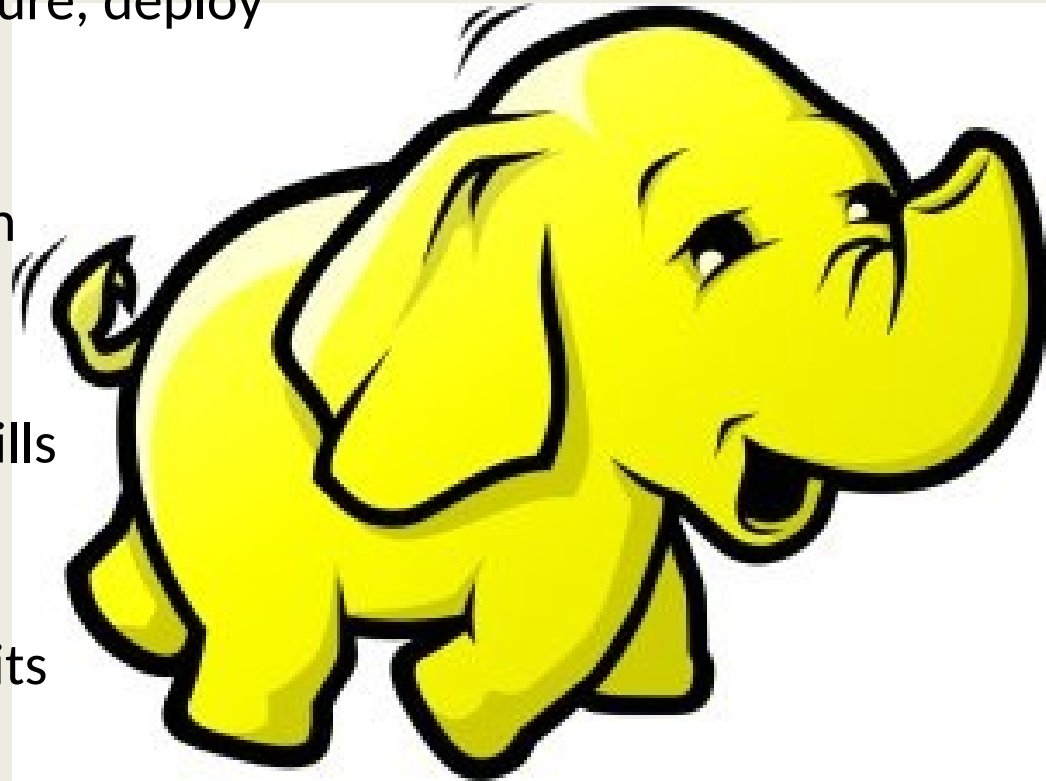
Source: [“Big data reconsidered: it's the economics, stupid.”](#) 451 Research, Dec 2013.

How to think strategically about big data

'Big Data' is the realization of competitive advantage by storing, processing and analyzing data that was previously ignored due to the cost and functional limitations of traditional data management technologies to handle its volume, velocity and variety

Zeroing in on Hadoop - barriers to Hadoop adoption

- Hadoop is complex to configure, deploy and manage
- Skilled staff are at a premium
- Enterprises want to make the most of existing tools/skills
- Enterprises are still trying to understand where Hadoop fits in their data management landscape



Your homework...

1. What business problem are you solving? What *questions* will you ask The Data?
2. Baseline existing costs, monitor new costs - did you save?
3. Monitoring and managing your new grid
4. Bonus: self-service access for *ad hoc* analysts

Thanks!

@cote – cote@451research.com – <http://cote.io>

