The Benefits of Modern Data Architecture

How Companies Can Slash Storage Costs, Eliminate Data

Silos, and Fully Realize the Promise of Big Data

A white paper from RCG Global Services

Executive Summary

_____

Traditional data architecture formats have served companies admirably for decades. But times, and data, have changed. Traditional data architecture formats can fulfill the needs of today's companies no more effectively than an ancient room-sized, vacuum-tube-powered UNIVAC computer.

Most companies are currently overwhelmed by the massive quantities of data available to them - a problem that's expanding at exponential rates. And much of the modern data available to companies is unstructured data from external sources, making it difficult or even impossible to effectively manage the data using traditional architecture formats.

At most companies the problem is further compounded by the proliferation of data silos. Within a single company, data may be replicated dozens, even hundreds of times across different business units. This replication of data renders the much-sought-after holy grail of an enterprise-wide "single version of truth" virtually unattainable.

The solution to these problems is the implementation of modern data architecture. This new architecture facilitates the storage and processing of both external and internal data in a single platform. It accommodates both structured and unstructured data. It's infinitely, easily, and affordably scalable to accommodate perpetually expanding needs. And it can eliminate data silos, creating a true enterprise-wide single version of truth.

Modern data architecture merges the traditional features and capabilities of an enterprise data warehouse with Hadoop. The result is a data management system that's far more efficient, effective, and much less expensive - a system that finally, _fully_ delivers upon the promise of big data.

The Data Deluge

As if they were engulfed by a digital tsunami, companies today are getting slammed with a massive tidal wave of data. But unlike a natural tsunami that eventually weakens and degrades into nothingness, the digital tsunami of data is expanding and strengthening at exponential rates.

Possible pull quote: Walmart collects more than 2.5 petabytes of data *hourly*.

Ninety percent of all the data that exists worldwide was generated in only the past few years. In 2012, approximately 2.5 exabytes of data was generated every single day. [1] Narrowed in scope to enterprise level, the deluge of data that companies are often faced with on a daily basis is still quite astounding. Walmart, for example, collects in excess of an estimated 2.5 petabytes of data on an hourly basis *just* from customer transactions. [2]

Possible pull quote: Most of the data generated today is unstructured, much of it from external sources.

And a major portion of this digital tsunami is comprised of unstructured data - data that often defies management using traditional data architecture systems. Roughly 75 percent of data generated today is unstructured in traditional terms.

Video, voice, and text are all examples of unstructured data that has proliferated wildly in just the past few years. Consider just one source of massive volumes of unstructured data: mobile phones. Already in use by more than half of the world's population, mobile phones will be used by nearly 70 percent of the global population by 2017. [3]

Data Silos Abound

The replication of data in multiple silos within an organization puts data integrity at risk. Siloed data works against the attainment of an enterprise-wide single version of truth. The mounting expense of storing all of that duplicated data also represents a massive waste in many organizations.

Possible pull quote: Data silos have become an ever-increasing problem.

But traditional data management systems have encouraged the proliferation of multiple data silos in most organizations. And like mushrooms springing to life after a long and soaking rainfall, the ever-increasing deluge of data has combined with traditional data architecture to foster an ideal environment for data silos. According to a recent Gartner report, it's likely that 90 percent of big data assets will be siloed by 2017. [4]

Companies Are Unprepared for the Challenge of Big Data

Many companies are quite unprepared and ill equipped to handle the ongoing massive influx of data. A 2014 report by Gartner predicts that "by 2017, 33 percent of Fortune 100 organizations will experience an information crisis." This crisis will be the result of organizations' "inability to effectively value, govern and trust their enterprise information." [5]

Possible pull quote: A third of the world's largest companies will experience an information crisis.

Even the simple expansion of data storage capacity can be a traumatic process with traditional systems. That's because traditional data management systems can only be scaled vertically. Once capacity is reached, the only option is to discard existing equipment and replace with equipment that offers greater storage capacity. Vertical scaling requires that all of the data residing on the old equipment be "forklifted" to the replacement equipment - an expensive, time-consuming, and risky process.

_____

Note: Place this sidebar wherever it best fits graphically anywhere from this point onward.

Sidebar: **Companies Using Modern Data Architecture**

Which companies have upgraded to a modern data architecture utilizing Hadoop®? You might recognize a few of these names… [7,8]

- Adobe
- Amazon
- AutoTrader
- Cablevision
- Ebay

The Solution: Modern Data Architecture

The solution to the data deluge is the implementation of a horizontally scalable system. Hadoop® is open-sourced software that facilitates the processing of data across clusters of servers. It can be horizontally scaled to any size necessary in accommodating the needs of an enterprise. Hadoop® systems may range in size from a single server to many thousands.

The advantages offered by Hadoop® includes:

- Accommodating both structured and unstructured data with unlimited scalability
- Handling the storage of data from internal and external sources
- Providing potentially huge cost savings over traditional systems
- Eliminating silos, and fostering a single version of truth
- Supporting a variety of work loads, including batch processing, real-time processing, interactive queries, and analytics

Using Hadoop® in conjunction with traditional relational database management systems provides companies with unprecedented capabilities in storing and managing data. The complimentary strengths of the two systems enable the management of all types of data, at unlimited volumes.

Conclusion: Reap the Benefits of Going Modern, Or Lose Ground to Competitors

Just like a natural tsunami, the digital tsunami of data that assails modern-day companies is wreaking havoc and causing chaos. That's because many enterprises that continue to rely upon traditional data architecture are simply incapable of handling the ever-increasing flood of data.

According to a 2015 Ernst & Young Report, *Becoming an analytics-driven organisation to create value*, more than 80% of business executives realize the importance of utilizing the data that's available for improving business efficiency and performance. But nearly a third admit to being overwhelmed by the sheer volume of available data. [6]

And yet there is a very distinct silver lining to this cloud. The wealth of data available to today's enterprises offers them unprecedented opportunities to more fully engage customers, increase business efficiencies, eliminate data silos, and achieve an enterprise-wide single version of truth.

The key to unlocking the true potential of big data lies in merging the capabilities of traditional data management systems with a horizontally scalable Hadoop®-based system. That's the *new* standard in modern data architecture.

Notes

1. Wall, Matthew. "Big Data: Are you ready for blast-off?" *bbc*. British Broadcasting Corporation, Mar. 2014. Web. 12 June 2015.

2. McAfee, Andrew, and Erik Brynjolfsson. "Big Data: The Management Revolution." *Harvard Business Review* Oct. 2012: n. pag. Web. 11 June 2015.

3. Wall, Matthew.

4. "Information Innovation Key Initiative Overview." *Gartner*. Gartner, Inc., 22 April 2014. Web. 11 June 2015.

5. "Gartner Says One Third of Fortune 100 Organizations Will Face an Information Crisis by 2017." *Gartner*. Gartner, Inc., 27 Feb. 2014. Web. 11 June 2015.

6. Ernst & Young LLP, and Nimbus Ninety. *Becoming an analytics-driven organisation to create value.* 2015. PDF file.

7. Henschen, Doug. "Hadoop 'No Longer Optional,' Says Forrester." *Information Week* Nov. 2014: n. pag. Web. 13 June 2015.

8. "PoweredBy." *wiki.apache*. Apache Hadoop, n.d. Web. 13 June 2015.