Big Data in Information Security

By Matt Sharp, Strategic Services Director

The ousting of Target CEO Gregg Steinhafel was at least partially due to a major data breach (Vigna, 2013). By observation, this has placed the topic of Information Security at the top of executive agendas throughout the country. At the same time, the Big Data discipline continues to mature such that 91% of executives claim that their organization has a Big Data initiative planned or in progress (Davenport T. H., 2013).

In this paper, we will explore how Big Data is different from data approaches of the past and attempt to extrapolate how Big Data will impact Information Security approaches today and in the future.

Data Analysis

To understand how Big Data is affecting Information Security, we first turn our attention to a short definition of Big Data followed by a summary of how businesses are using big data. Finally, we will examine the impact of such technologies on Information Security.

What is Big Data, and how is it different from previous approaches?

Big Data is a term used to describe collections of data so large, complex or requiring such rapid processing (sometimes called the volume/variety/velocity problem), that they become difficult or impossible to work with using standard database management or analytical solutions (Davenport T. H., 2013).

Big Data is frequently used to quantify the amount of data in the literal sense; however, most experts "in the know" consider this to be a marketing term. The true analytics and knowledge that power such initiatives are more commonly referred to as "data science."

Nevertheless, I will stick with the term Big Data for this briefing. Organizations that capitalize on Big Data stand apart from traditional data analysis environments in three key ways (Davenport, Barth, & Bean, 2012):

Abstract

As the Big Data discipline continues to mature, the impact to information security will be felt. This paper examines Big Data, how businesses are using Big Data and how information security approaches will need to adapt.

Table of Contents:

- They pay attention to data flows as opposed to static data warehouses.
- They rely on data scientists as well as product and process developers rather than data analysts.
- They are moving analytics away from the IT function and into core business, operational and production functions.

Business Value of Big Data

"Investments in Big Data include those in human resources and technology solutions, including database management platforms (e.g., Hadoop, EMC/Greenplum, Teradata/Aster, IBM/Netezza), analytics and visualization capabilities (e.g., Revolution R, Palintir, Tableau), or text-processing and real-time streaming solutions" (Davenport T. H., 2013). *See Figure 1.*

Interestingly, most Big Data initiatives do **not** currently require an ROI payback analysis to justify their current investment. However, analysis shows that executives were expecting three primary benefits resulting from Big Data investments including (Davenport T. H., 2013):

- Accelerate the analytical velocity of delivering insights and answers to business questions.
- Integrate a greater variety of data sources.
- Analyze larger volumes of data.

Surprisingly, **most** Big Data initiatives were linked to efforts that enhance analytics capabilities (66%) and enterprise risk and analytics initiatives (54%)" (Davenport T. H., 2013). *See Figure 2.*

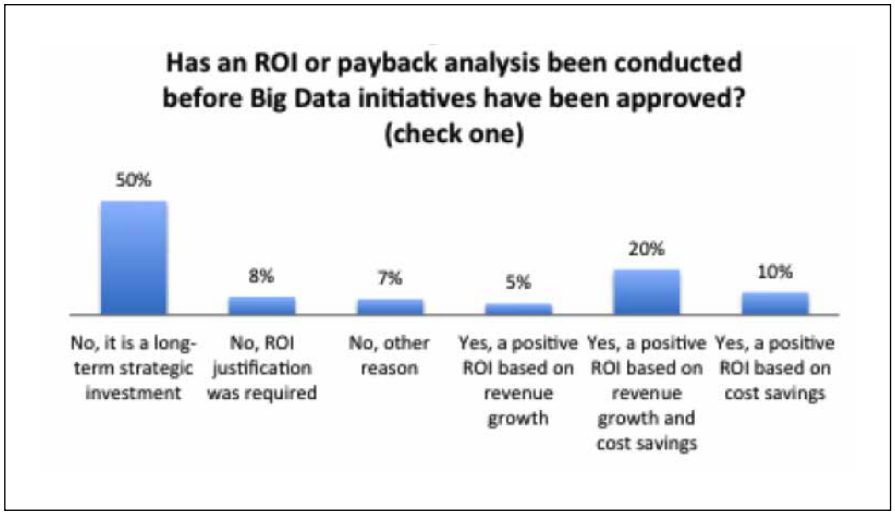

Figure 1

Big Data Impact on Information Security

Big Data and Information Security overlap to create several distinct outcomes. Big Data improves the speed at which firms can identify and respond to threats, and at the same time, Big Data can have significant privacy implications as the aggregated data is a new target for hackers.

Accelerated Threat Detection

In short, firms are largely utilizing investments in Big Data to help security teams detect threats quicker and speed up response. One notable study stated:

"Critical to controlling the costs [of a data breach] is keeping customers from leaving. The research reveals that reputation and the loss of customer loyalty does the most damage to the bottom line. In the aftermath of a breach, companies find they must spend heavily to regain their brand image and acquire new customers... Efficient response to a breach and containment of the damage has been shown to reduce the cost of breach significantly" (Ponemon Institute, 2014).

ID#6LABS.WP0006 Last Modified 06.03.2014

Figure 2

Indeed, "threat analysts need a combination of capture time, stream and batch analytics to detect and investigate a full range of threats" (EMC Corporation, 2014). "There [are] so many events happening at the network layer, the ability to do stream processing across those events and detect anomalous, malicious behavior is important" (Prince, 2014).

Logically, this is a natural result of the fact that "IT systems and technology infrastructure generate data every second of every day. This machine data contains a categorical record of all user behaviors, service levels, cybersecurity risks, fraudulent activities and more. As one of the fastest growing and most complex segments of big data, machine data is also one of the most valuable" (Splunk, Inc., 2014).

Expanded Attack Surface

However, Big Data is also a target for hackers. In fact, "just as hackers find a clever exploit to intercept and spike an SSL session, or trick an app server into running arbitrary code, so they'll find an exploit for big data… Attackers can corrupt information, blind an algorithm, inject falsehood, changing outcomes in subtle, insidious ways that undermine a competitor or flip an election. Attacks on data become attacks on people" (Webb & O'Brien, 2013).

Finally, it is important to note that "privacy protection has become an elusive goal in the big data era as researchers have shown that 'linkability threats' can re-identity individuals… In practice, such data is shared after sufficient removal of unique identifiers by the processes of anonymization and aggregation. This process which has led to very many instances of re-identification based on big data linkability needs to be strengthened." (Cloud Security Alliance, 2014).

ID#6LABS.WP0006 Last Modified 06.03.2014

Conclusions

It turns out that volume, variety and velocity challenges addressed by Big Data are incredibly relevant to the analysis of anomalous and malicious activity on enterprise networks. While recent industry events have elevated the topic of Information Security in the board room and investments in Big Data continue to rise, the overlap of Big Data and Information Security present significant business opportunities.

Exploiting the overlap of these emerging trends to reduce corporate risk exposure through rapid identification of enterprise threats is likely to produce more indirect benefits and therefore can be very difficult to justify with ROI calculations.

Nevertheless, the consensus seems to be that "modern risk management requires real-time data and business self-sufficiency so risk owners can respond to business, board and regulator demands in a timely and accurate fashion" (Prince, 2014).

At the same time, it is important to understand the security and privacy implications resulting from Big Data implementations supporting non-Information Security functions. Specifically, security obligated executives should be cognizant of how Big Data increases attack surface for hackers and understand how to protect against linkability threats.

For More Information

For more information about FishNet Security products and services, call 888.732.9406 or visit: *www.fishnetsecurity.com.*

Works Cited

Cloud Security Alliance. (2014). *Comment on Big Data and the Future of Privacy.*

Davenport, T. H. (2013). Big Data Executive Survey 2013: The State of Big Data in the Large Corporate World. Boston | San Francisco | New York: NewVantage Partnes LLC.

Davenport, T., Barth, P., & Bean, R. (2012). How 'Big Data' is Different. MIT Sloat Management Review.

EMC Corporation. (2014). RSA-Pivotal Security Big Data Reference Architecture. EMC Corporation.

Ponemon Institute. (2014, March 05). 2014 Cost of Data Breach: Global Analysis. Retrieved from Ponemon Institute: http://www.ponemon.org/blog/ponemon-institute-releases-2014-cost-of-data-breach-global-analysis

Prince, B. (2014, February 27). Big Data A Big Focus Of Security Analytics Products. Retrieved from Dark Reading: http://www.darkreading.com/big-data-a-big-focus-of-security-analytics-products/d/d-id/1141401?

Splunk, Inc. (2014, May 10). Big Data Analytics: Delivering Insights on Real-time and Historical Data. Retrieved from Splunk : http://www.splunk.com/view/big-data/SP-CAAAGFH

Vigna, P. (2013, May 05). Retrieved from The Wall Street Journal: http://blogs.wsj.com/moneybeat/2014/05/05/targets-ceo-ouster-wasnt-about-just-the-data-breach

Webb, J., & O'Brien, T. (2013). Big Data Now. O'Reilly Media, Inc.

ID#6LABS.WP0006 Last Modified 06.03.2014