

Big Data Options For Small And Medium Enterprises

Donna M. Schaeffer, Marymount University, USA
Patrick C. Olson, National University, USA

ABSTRACT

The terms Cloud Computing and Big Data are often used in the media. Small and medium sized businesses are realizing that the amount of data they handle could be very large and important in their decision making and planning. This paper explores the options for handling large amounts of, or “Big” data in small and medium-sized enterprises. This research considers not just the volume of data, but the variety of types of data, the velocity in which data is created and transmitted, as well as the importance of data veracity. Cloud computing is a potential solution.

Keywords: Big Data; Cloud Computing; Small and Medium Business Enterprises

INTRODUCTION

The United States Small Business Administration (2012) (SBA) defines a small business concern as one that is independently owned and operated, is organized for profit, and is not dominant in its field. Depending on the industry, size standard eligibility is based on the average number of employees for the preceding twelve months or on sales volume averaged over a three-year period. Examples of SBA general size standards include a maximum number of employees ranging from 500 to 1500, depending on the type of product manufactured or from 100 to 500 employees depending on the particular product in a wholesaling enterprise. Annual receipts may not exceed \$2.5 to \$21.5 million, depending on the particular service being provided; or \$5.0 to \$21.0 million in retail enterprises. In general and heavy construction, annual receipts may not exceed \$13.5 to \$17 million, depending on the type of construction; or in special trades, annual receipts may not exceed \$7 million. In agriculture, annual receipts may not exceed \$0.5 to \$9.0 million, depending on the agricultural product.

Although large corporations garner most media attention, there are some indications that small businesses are a major player in the United States' economy. Small businesses represent more than 99 percent of U.S. firms with employees and account for 49 percent of private sector employment (SBA, 2012). Freudenberg et al. (2012) indicated that small businesses account for 43% of sales and 51% of the Gross Domestic Product in the United States. The Small Business Administration Office of Advocacy (2012) estimates that, of the 18.5 million net new jobs created in the United States between 1993 and 2011, small businesses (less than 500 employees) accounted for 11.8 million, or 64 percent.

SMALL BUSINESS AND INFORMATION TECHNOLOGY

Given the wide variety of industries and locations which small businesses operate, as well as the differences in numbers of employees and amounts of revenue, it is understandable that the adoption of information technology (IT) in small businesses varies.

The National Small Business Association (2010) found that computers have a significant role in small business with 98% of respondents of a survey agreeing that it is important to keep up with technology. Many respondents reported having websites, purchasing supplies online, banking online, and using social media for business networking. Kim et al. (2013) studied the Web 2.0 practices of one hundred small and medium sized enterprises (SMEs) in the United States. The organizations consisted of those ranked one of the fifty best SMEs to work for in America, as well as fifty randomly selected enterprises. Most of the fifty best SMEs had adopted Web

2.0 in various degrees, but the other SMEs needed to increase their efforts to improve their performances, to connect with consumers and to remain competitive.

IDC (2012) reported spending on information technology by SMEs in the United States to be \$53.0 billion in 2011 and expected that figure to exceed \$138 billion in 2012. SMEs account for 10% of the investment in information technology worldwide (IDC, 2012). SMEs are spending increasing more of their budgets on cloud technology (Dyson, 2013). As a result, spending on security is increasing and is expected to pass \$5.6 billion in 2015 (Eddy, 2012). A survey commissioned by Brother International Corporation (2013) also indicates just over half of small business owners surveyed (51%) viewed technology tool-related capital investments such as new software, mobile apps, and cloud computing services as a top priority for investment.

A survey (Kass, 2013) of 173 small business owners or managers users demonstrated they are actively using five different devices including laptops, desktops, smartphones, and tablets. Some owners and managers reported regularly using as many as 12 different devices (Kass, 2013).

THE DIMENSIONS OF BIG DATA

“Big data” is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making (Beyer & Laney, 2012). In addition to considerations about its volume, velocity, and variety, big data requires veracity.

High volume refers to data sets with sizes ranging from a few dozen terabytes to many petabytes of data in a single data set. This volume of data is typically beyond the ability of commonly used software tools to capture, curate, and store data (Kuznetsky, 2010) as well as the ability to process the data within a tolerable elapsed time (Snijders, Matzat, & Reips, 2012). The volume also complicates the ability to share, transfer, and analyze the data (Vance, 2012).

While conducting a large project, data processing and storage needs could rise to the petabyte level. A petabyte is a term for an amount of computer storage; it is a unit of information equivalent to 1000 terabytes or 10^{15} bytes. While terms like bit and gig have entered everyday vocabulary, authors try to explain a petabyte to the general public. A petabyte has been described as four times the amount of information in the United States Library of Congress or approximately 223,000 DVDs (McKenna, 2013), 20.0 million four-drawer filing cabinets filled with text (Mozy, 2009), or the memories of 800 human beings (Kurzweil, 2005). The amount of information generated is expected to grow fifty times over by 2020 (Nevitt, 2013). IBM (2012) has estimated 2.5 quintillion bytes of data are created daily. Frank Moss (2011) stated:

Computers, smart phones, GPS devices, embedded microprocessors, sensors — all connected by the mobile Internet — are forming a ‘societal nervous system’ that is generating a cloud of data about people that is growing at an exponential rate. Every time we perform a search, tweet, send an email, post a blog, comment on one, use a cell phone, shop online, update our profile on a social networking site, use a credit card, or even go to the gym, we leave behind a mountain of data, a digital footprint, that provides a treasure trove of information about our lifestyles, financial activities, health habits, social interactions, and much more.

The velocity aspect of big data required high speed connectivity and broad bandwidth. Small business owners in large urban areas may have sufficient bandwidth to handle big data projects, but those in the more rural areas of the United States may not have access to high-speed bandwidth. Drobo's (2012) study discovered 46 percent of small business owners don't have the bandwidth or sufficiently reliable networks for cloud storage. Rather, small business owners opt for a hybrid backup and storage solution of onsite and offsite.

Big data is inherently high variety; but the data generated and used by small businesses may be more limited in scope. The limited variety of data inhibits the predictive capabilities and discovering the unexpected information that big data projects tout as beneficial. Mowbrey (2009) describes four different types of data that a small business might typically store in a cloud. The four types of data include information about customers, accounts such as payments, operating data, and activity data that tracks when and for which applications the business' account with the service provider is used.

THE SYSTEM OF BIG DATA

To effectively gather and use big data for decision making, a system is required to capture, curate, manage, and process the data. Systems, or large parts of systems, may already be in use in a small business. For example, a small business with a web presence can track online visitors to its site via a simple technology such as cookies or more sophisticated visitor analytic tools. Small business owners can track who is visiting the site, where they are coming from, how long they stay, and which pages they browse. This information can be used in targeted marketing campaigns.

Data can be captured via point-of-sale (POS) terminals or other devices. POS terminals are the hardware used to manage the selling process with a salesperson-accessible interface in the location where a transaction takes place. During 2012, within the United States, there were 2,156 POS terminals per 100,000 adults (World Bank). GAO Research, Inc. (2012) estimates there are 10 million payment terminals in the United States, but there is no data about how many of these systems are in place in small businesses. With advances in Near-Field Communications (NFC) technologies, POS terminals are evolving into mobile payment systems. NFC is a set of standards that enable wireless communications between smartphones that are in very close proximity.

Canicola (2013) identified several technologies that small businesses may have in place via which data can be captured – customer relationship management (CRM) systems, websites and visitor/lead analytic tools, email marketing tools, blogs, or electronic payments. Data could also be captured via web-based Voice of Internet Protocol (VoIP) telephone systems (Totka, 2013). In many cases, these technologies are provided to small businesses by other firms, such as Salesforce.com for CRM or Google for analytics, or tools like Constant Contact or Wordpress for specific applications.

The volume of data that is available due to POS, mobile payment systems, or other systems can be overwhelming to a small business owner. As mentioned earlier, data can quickly reach the petabyte level. Many small businesses opt to store their data using services in “the cloud.” For example, Amazon Web Services provides the use of Amazon’s computing infrastructure, including a power, storage, and e-commerce software. Other examples include Google Apps, a platform on which companies they can run and store their own applications and data in the cloud and Salesforce.com, a customer relationship management software.

BIG DATA, SMALL BUSINESS, AND THE CLOUD

The use of the cloud for data storage among small businesses is mixed. The Brother Small Business Survey (2013) reports that 35% of respondents use the Cloud for data storage, but 42% are not using the cloud at all. Perhaps small business owners don’t understand the concept – 46% of respondents understood “cloud” somewhat, while 27% responded that they didn’t understand it well or at all. This is consistent with other findings. A survey of 239 small and medium sized enterprises found that 46% of respondents believe that public cloud storage is not a fit for their business (Drobo, 2012).

Mowbrey (2009) points out that Service Level Agreements (SLAs) for cloud computing services do not always make it clear what rights the providers have to use the data they store. For many providers, the data is only monetarized when it is sold. Small business owners need to seek clarification. For example, the Amazon S3 storage service has “buckets” around the world, and users may choose which the region in which they want to store their data. Choice is factored by issues such as geographic proximity to handle latency issues, geographic remoteness for redundancy and disaster recovery purposes, legal and regulatory requirements, or costs.

The use of the cloud raises data veracity issues. A study conducted by Techaisle (2012) shows that it is important for small and medium enterprises to maintain ownership of data and that may not always be the case with cloud services. In some industries, there are strong regulations about data; for example, the Attorney General in Minnesota filed suit with a health care provider over lax security of thousands of patients’ digital health records (Lynn, 2013).

CONCLUSION

Considering the needs noted above, small and medium businesses need to work with big data. All indicators show that the volume of data is rising and will continue to do so simply because of the trends related to customer interaction. This leaves small and medium businesses with the problem of how to participate. The cloud is a good option. However, several sources indicate SMEs are reluctant to enter the cloud. This is likely related to issues of control, particularly ownership. Additionally, there is an access problem due to the difference in the way small and medium businesses access the Internet. Realistically, the difference isn't that surprising since many smaller companies still rely on consumer broadband and wireless services for access. Only one-third of small and medium sized businesses have actually invested in networking infrastructure, as a recent IDC (2012) study estimated.

One of the important benefits of big data is the ability to discover relationships that were not previously known, but this benefit may be unavailable to SMEs. This is likely because while small and medium businesses will certainly generate the volume of data that is associated with big data, they are unlikely to generate the variety of data usually associated with big data. Interestingly, a collection of SMEs would have that variety. However, accomplishing this would require the technology and the bandwidth to carry it.

Because access to bandwidth is a problem for more than one aspect of SME's involvement in big data, might this be a time to expand the opportunities provided by the Internet? The Internet2 Consortium recently upgraded its network backbone to a 100-Gigabit, Ethernet-enabled, 8.8-terabit-per-second optical network. Though usually used by education and government, allowing access to this network by the SME's would increase available bandwidth.

AUTHOR INFORMATION

Dr. Schaeffer has taught at universities in the United States, Germany, and Korea. She earned her PhD in the Management of Information Systems at Claremont Graduate School. Dr. Schaeffer is a participant on the Internet Advisory Caucus for the U.S. Congress and the Business Advisory Forum of the United Nations. She has published over 50 articles and book chapters. She is active in the Churchill Club, Computer History Museum, Consortium of Computing Sciences in Colleges, Decision Sciences Institute, John Dewy Society, Institute of Electrical and Electronics Engineers, Women's High Tech Forum, Women in International Security, and World Affairs Council. E-mail: dschaeff@marymount.edu (Corresponding author)

Dr. Olson is a Professor at National University. He earned his PhD in the Management of Information Systems at Claremont Graduate School. He also holds a Master of Science in Systems Management from the University of Southern California and Bachelor of Arts in Communications from University of Montana. He worked in industry at Hughes Aircraft and held administrative and faculty positions in higher education since 1988. Dr. Olson has served as an Examiner for the State of California Malcolm Baldrige Quality Awards. E-mail: polson@nu.edu

REFERENCES

1. Beyer, M. A., & Laney, D. (2012) *The importance of 'big data': A definition*. Gartner Report, June version, ID G00235055.
2. Brother International Corporation. (2013, March 19). *Brother small business survey 2013*. Retrieved from <http://www.brother-usa.com/PressReleases/brother%202013%20smb%20survey%20press%20release%20-%20final.pdf>
3. Cenicola, M. (2013, April 29). Three steps to incorporate big data into your small business. *Forbes*. Retrieved from <http://www.forbes.com/sites/theyec/2013/04/29/3-steps-to-incorporate-big-data-into-your-small-business/>
4. Curry, E., Freitas, A., & O'Riain, S. (2010). The role of community-driven data curation for enterprises. In D. Wood (Ed.), *Linking enterprise data* (pp. 25-47). Boston, MA: Springer.
5. Drobo Incorporated. (2012). *SMB IT 2012 Survey Results*. Retrieved from http://www.drobo.com/downloads/docs/SR-0137-00_SMB-virt-survey-q3-q4-2012.pdf

6. Dyson, C. K. (2013, January 9). Can the cloud help small businesses? *Wall Street Journal*. Retrieved from <http://online.wsj.com/article/SB10001424127887323706704578230641145851624.html>
7. Eddy, N. (2012, May 30). *Small-business security spending to top \$5.6B in 2015*. Retrieved from <http://www.eweek.com/c/a/Security/Small-Business-Security-Spending-to-Top-56-Billion-in-2015-IDC-875794/?kc=rss>
8. Freudenberg, B., Tran-Nam, B., Karlinsky, S., & Gupta, R., (2012). A comparative analysis of tax advisers' perception of small business tax law complexity: United States, Australia and New Zealand. *Australian Tax Forum*, 27(4), 677-718. Retrieved from <http://ssrn.com/abstract=2190692>
9. GAO Research. (2012). Retrieved from <http://www.gaoresearch.com/POS/pos.php>
10. IDC. (2012, March) U.S. small and medium-sized business 2012–2016 forecast: Sizing the SMB markets for PCs and peripherals, systems and storage, networking equipment, packaged software, and IT services.
11. Kass, L. (2013, May 07). *Survey results reveal top five reasons small businesses are turning to cloud file management*. Retrieved from <http://www.sugarsync.com/blog/2013/05/07/survey-results-reveal-top-five-reasons-small-business-are-turning-to-cloud-file-management/>
12. Kim, H. D., Lee, I., & Lee, C. K. (2013). Building Web 2.0 enterprises: A study of small and medium enterprises in the United States. *International Small Business Journal*, 31(2), 156-174.
13. Kurzweil, R. (2005) *The singularity is near*. New York: Viking.
14. Kuznetsky, D. (2010). "What is big data?"
15. Lynn, S. (2013, June 4). Small business cloud myths: Busted!" *PC Magazine*. Retrieved from <http://www.pcmag.com/article2/0,2817,2419823,00.asp>
16. McKenna, B. (2013). What does a petabyte look like? *Computer Weekly*. Retrieved from <http://www.computerweekly.com/feature/What-does-a-petabyte-look-like>
17. Mowbray, M. (2009). *The fog over the grimpen mire: Cloud computing and the law*. HP Laboratories: HPL-2009-99. Retrieved from <http://www.hpl.hp.com/techreports/2009/HPL-2009-99.pdf>
18. Moss, F. (2011). *The sorcerers and their apprentices: How the digital magicians of the MIT media lab are creating the innovative technologies that will transform our lives*. New York: The Crown Publishing Group.
19. Mozy, Inc. (2009). How much is a petabyte? Retrieved from <http://mozy.com/blog/misc/how-much-is-a-petabyte/>
20. National Small Business Association. (2010). *Small business technology survey*. Washington, DC.
21. Nevitt, C. (2013, March 18). What is a petabyte? *Financial Times*. Retrieved from <http://www.ft.com/cms/s/2/bc7350a6-8fe7-11e2-ae9e-00144feabdc0.html#axzz2WD55Twcz>
22. Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., & Tufano, P. (2012). *Analytics: The real-world use of big data*. Somers, NY: IBM Corporation.
23. Snijders, C., Matzat, U., & Reips, U.-D. (2012). Big data: Big gaps of knowledge in the field of internet science. *International Journal of Internet Science*, 1, 1-5. Retrieved from http://www.ijis.net/ijis7_1/ijis7_1_editorial.pdf
24. Sugarsynch. (2013). Retrieved from <http://www.sugarsync.com/blog/2013/05/07/survey-results-reveal-top-five-reasons-small-business-are-turning-to-cloud-file-management/>
25. Totka, M. (2013, May 21). How a small business can use big data. *Small Business Operations*. Retrieved from <http://smallbiztrends.com/2013/05/small-business-can-use-big-data.html>
26. United States Small Business Administration. (2012).
27. Vance, A. (2010, April 22). Start-up goes after big data with hadoop helper. *New York Times Blog*. Retrieved from <http://bits.blogs.nytimes.com/2010/04/22/start-up-goes-after-big-data-with-hadoop-helper/?dbk>
28. World Bank. (2012). Retrieved from <http://data.worldbank.org/indicator/FB.POS.TOTL.P5>

NOTES