

Big Data Survey

Real-Time Stream Processing and Cloud-Based Big Data Increasing in Today's Enterprises

Companies, Increasingly Seeing Big Data as Mission-Critical, Want Two Things: To Take Big Data to the Cloud, and Be Able to Process Real-Time Streaming Events.

Executive Summary

Everyone has been speaking about Big Data for some time, but there is a lot of speculation and tire kicking by many in the industry. We wanted to evaluate how people are viewing Big Data today, how important it is, how to deal with it, what returns can be had from it... Aside from rampant speculation, what are companies actually doing and planning regarding Big Data *today*? Do they see a need to adopt new technologies? What direction are they going in? What are their major concerns?

In Fall 2012, GigaSpaces conducted a survey to find out more about what IT professionals are concerned with regarding their Big Data systems, what they are planning, and what tools they are using.

The survey resulted in three major findings:

- An overwhelming majority of organizations view their Big Data processing as mission-critical.
- For companies handling Big Data, the need for real-time functionality is both significant and growing. The survey indicated that there is increasing readiness to use streaming solutions to deal with the challenges of Big Data and speed up Big Data processing.
- Most companies have plans to move their Big Data to the Cloud, or are considering the option.

Only 20% of the IT professionals surveyed indicated that their company had no plans to move their Big Data to the cloud.

Notes on Methodology

The survey was conducted via an online survey service aimed at IT audiences, and distributed to IT and business professionals at several industry trade shows. There were a total of 243 respondents, including developers, IT managers, architects, and C-level executives, from 8 different countries across North America, Latin America, and Europe/UK. Companies surveyed belong to a variety of industries, including Financial Services, Telecom, Retail, Insurance, Media, Government, E-Learning, Software Development, SaaS, and more.

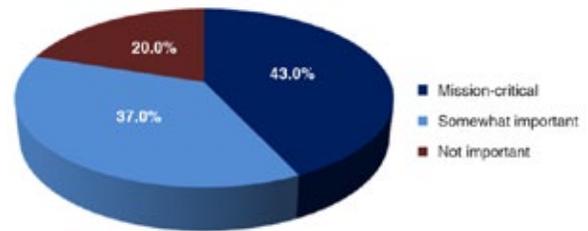
Results

Importance of Big Data

The first goal of our survey was to assess how important Big Data is to companies, today. 80% of responders indicated that Big Data was important to their business, with 43% indicating it is mission critical.

Responder comments included indications that data traffic is rapidly increasing, that processing the data is of mission-critical importance, and that back-up, querying, and security were of high importance in their businesses' Big Data processing.

How important is Big Data processing to your business?

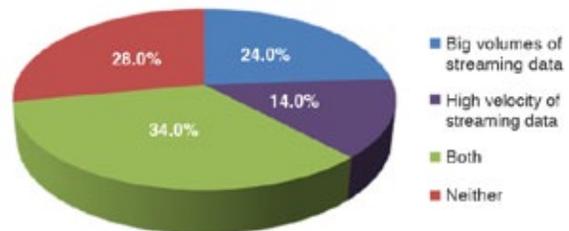


The Need for Real-Time Processing

GigaSpaces wanted to also determine what kind of processing was necessary for the responding companies' Big Data. Over 70% of responders indicated that they process streaming Big Data – either at large volumes, high velocity, or both. Only 28% said that they do not processing streaming events.

Some responders commented that though they are not streaming Big Data today, they are planning to soon, or are looking into new ways to gather and process data. Today, they are using various methods to do so.

Do you process streaming Big Data in your business

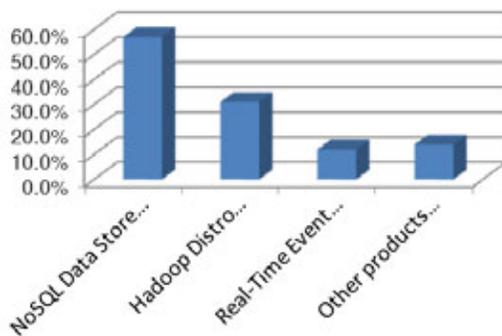


The need to process *high velocity* of Big Data relates to the need for *low latency*, such as is necessary for automatic trading, telecom fraud detection, or e-commerce applications. The need to process *large volumes* of data refers to *performance*, or the need to ensure that the speed of data entry and data processing are aligned so that no backlogs are created, as is necessary for analytics apps, for example.

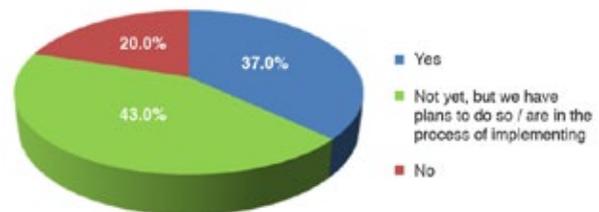
What are the Tools for Big Data?

A whopping 80% of responders said they are already using or planning to use dedicated Big Data tools or architectures in their production environment to cope with the influx of massive amounts of data. 56% of respondents indicated they plan to move existing applications from RDBMS to a NoSQL data store.

If so, which of the following products?



Are you using any tools/architecture in your production environment to cope with Big Data?



Over 30% of those currently using dedicated tools are using a combination of products, such as a NoSQL data store together with a Hadoop processing platform. Interestingly enough, only 12% are utilizing real-time event processing tools, in spite of the 70% who indicated a need for stream processing.

These findings suggest a possible gap in the market, where enterprises still have not found the right solutions that combines the ability to handle massive data while also providing speed. They also confirm a market trend some have already noted, whereas organizations requiring both extreme performance and scalability are moving beyond Hadoop (or at least Hadoop-only) architectures to handle Big Data systems.

Some responders commented while they are working with NoSQL data stores, such as Cassandra, they only route to the NoSQL store data not requiring transactionality, while any processes requiring transactionality are dealt with via RDBMS and other legacy architecture. This clearly underscores the need for solutions that can work with data distributed at extreme scale, while providing transactionality and maintaining both read and write consistency.

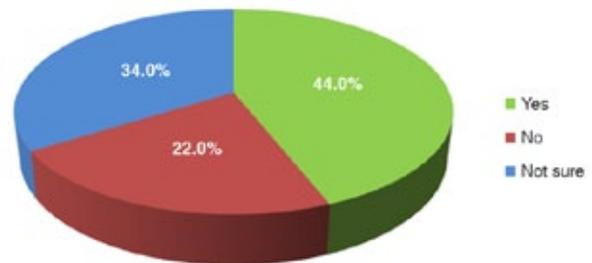
Big Data on the Cloud?

As Big Data invades the enterprise, there is also an increasing desire to use the cloud to gain benefits of economy and agility. The survey also aimed to find out at what point these two areas met: Is the cloud the next step for Big Data applications?

The results were somewhat surprising in the degree to which these areas intersect – only 20% of respondents indicated they have no plans to move Big Data apps to the cloud. 44% of responders indicated they plan to move their Big Data application to the cloud or have already begun to do so. Among the 34% that were unsure, primary concerns cited regarding a move to the cloud were scalability and security. Some responders reported that they are working on a private cloud environment in order to gain the benefits of the cloud without the perceived or actual risks of working in a public cloud environment.

To get an even better idea of whether those moving or considering a move to the cloud are running mission-critical or non-critical Big Data apps, the questions of the business importance of Big Data was cross-referenced with the cloud plans question, and the results support the idea that enterprises want, plan, or need to move their business-critical apps to the public cloud: 80% of those defining their Big Data apps as “Mission-Critical” to the business are planning or considering a move to the cloud, while of those who consider their Big Data “Somewhat Important” the number was 75% for the cloud move.

Do you have plans to use the cloud for Big Data?



Conclusions

The survey clearly indicates that enterprises deem Big Data as essential to their business, that they are increasingly reliant on their ability to process Big Data effectively, and that they largely need that data processed in real time, in order to be able to make business decisions based on analysis of the data, or to handle continual data inputs.

It can be concluded from the findings that enterprises are still looking for the right infrastructure tools that will enable them to effectively handle their Big Data, in line with business needs. Most companies are already using dedicated Big Data tools, but all still see gaps in capabilities or have concerns regarding the fit between these tools and their current and expected needs.

Finally, the convergence of Big Data and cloud computing cannot be ignored: companies cannot get the full benefits of using the public clouds unless they are able to run their Big Data applications there as well, which in turn means that tools that guarantee consistency, transactionality, reliability, scalability, and security are crucial to their IT strategy.

About GigaSpaces Technologies

GigaSpaces Technologies is the pioneer of a new generation of application virtualization platforms and a leading provider of end-to-end scaling solutions for distributed, mission-critical application environments, and cloud enabling technologies. GigaSpaces is the only platform on the market that offers truly silo-free architecture, along with operational agility and openness, delivering enhanced efficiency, extreme performance and always-on availability. The GigaSpaces solutions are designed from the ground up to run on any cloud environment – private, public, or hybrid – and offer a pain-free, evolutionary path to meet tomorrow's IT challenges.

For more information, visit <http://www.gigaspaces.com>