BIG TRUCKS DELIVER BIG DATA AT NAVISTAR

Navistar, Inc. is a completely different company from the one that was founded in 1831 to manufacture McCormick farming machines and other agricultural equipment. One thing that hasn't changed,

however, is our focus on innovation. Our products helped drive the westward expansion in the U.S. with farming equipment and, over time, construction equipment, gas turbines, trucks, and buses. We like to think of ourselves as constantly "pushing the limits of what's possible and driving history forward."

The new frontier today is big data and the Internet of Things (IoT). We're leveraging big-data and IoT technologies in multiple areas to improve our products and create new offerings that deliver incredible value to Navistar customers. Our innovation initiatives require the movement, processing, and analysis of huge volumes of data. For example, we run a number of Sqoop™ jobs that gather data from internal backend database systems—Teradata, SQL server, DB2, Informix, and Oracle. We collect, process, and normalize data collected from external telematics providers, combine it with the internal data, bring it into Hadoop, and then do additional processing using a combination of Hive™ jobs and Python scripts.

Managing the workflow for these complex operations is challenging. In the past, the people working on big-data projects were spending a lot of time manually moving data and running various kinds of scripts. It was time consuming and it sometimes meant they didn't have access to actionable data for days. They needed a faster, more efficient way to handle big data.

We found that Control-M from BMC was exactly what we needed to streamline and automate batch processing for our big-data projects. For 17 years, Control-M has played a vital role in running Navistar's business. We use it to update databases, extract data and feed it to other systems, maintain price masters and parts masters, supply parts information to distribution centers, integrate with warehouse management systems for inventory and restocking, and handle financial reporting. We run close to 20,000 Control-M job executions daily across all our platforms, totaling nearly 3.3 million executions monthly.

Hadoop is an ecosystem in itself and has its own computing platform. The opportunities it presents are vast and we are leveraging Control-M to tap that power. Because Control-M is easy to use, we are transforming our scheduling processes by giving the applications teams more control over their scheduled processes. This is freeing up time for our schedulers so they can focus on ensuring adherence to standards and best practices.

Our Hadoop team is using Control-M to set up and schedule job streams. It's the first time we've allowed application developers to define their own jobs. They are creating jobs in a nonproduction environment and running the jobs manually. This allows them to change their configuration profiles or connection profiles to their platform and create profiles based on the data they're ingesting into Hadoop. We then assist the developers in migrating the jobs to production. The Hadoop team has already defined nearly 200 jobs.

Control-M is delivering value in such high-profile projects as Navistar's OnCommand™ Connection remote diagnostics system and quality improvement in truck design. These projects are beginning to deliver significant value to Navistar's business, enabling us to build higher-quality vehicles and ensure that customers get maximum value from the products they buy.

OnCommand™ Connection

We launched OnCommand Connection in 2013 and today we're using it to monitor the performance of more than 160,000 trucks. Every day we capture about 100 data points from each vehicle, including fuel economy, geographic location, idle time, and potential failures. OnCommand allows

us to monitor entire fleets, so we can diagnose issues as they develop and provide guidance to fleet owners—for example, helping them route trucks to dealer service centers that have the required parts readily available for quick servicing. For customers, this means less time in the garage and more time on the road, and that translates into a higher return on the investment in the vehicle. OnCommand can help customers reduce unplanned repairs and downtime by as much as 30 percent. In the future, we hope to monetize this diagnostic analytics capability and sell it as a service to smaller truck fleets.

To provide OnCommand services, we capture millions of data points every day and transform them into an actionable form. We collect data from 13 telematics providers, process and normalize it, augment it with data from internal systems, and deliver the result to Hadoop. Once the data is in Hadoop, Navistar engineers and data scientists can leverage it to build predictive models to understand when a truck needs service.

Control-M orchestrates this enormously complex movement and processing of huge amounts of data. With Control-M, we're confident that jobs are being successfully completed at the right time and in the right order. If a problem arises, Control-M immediately notifies the appropriate people. We have a lot fewer sleepless nights knowing Control-M is there.

Improving Commercial Truck Quality

We're putting the data we're collecting through OnCommand to work in other ways as well. Our engineers are using it to help identify design improvements that will make our commercial trucks safer, more efficient, and more reliable. Better quality means a competitive edge in a competitive market.

Another big-data initiative complements the OnCommand data with additional detailed data collected through special devices we've installed on a fleet of test trucks. Each device captures more than 12,000 data points. Our design engineers leverage this data to continually improve truck design.

We are partnering with a company that collects the data from approximately 60 trucks and stores it on Amazon Web Services (AWS). Control-M pulls the data from AWS and puts it in Hadoop. We tap into the data with queries built by the engineers. Ultimately, they use Tableau to process and display the information on dashboards. Prior to bringing Control-M into our Hadoop environment, we had two engineers working full time pulling this data from AWS, aggregating it, and putting it into a spreadsheet for the design engineers. It took about a week to get the data into a usable form. Control-M now handles these tasks automatically, freeing up those two engineers to take on strategic tasks. In addition, the design engineers now have real-time access to data that's viewable on easy-to-read dashboards, a marked improvement over static reports based on data that's a week old.

We're looking into other ways to use this data as well. For example, we foresee monitoring customer vehicles to get a handle on how driver habits affect maintenance and warranties. If we observe bad driving behaviors that result in a higher number of warranty repairs, Navistar can rethink warranty commitments and pricing and have meaningful conversations with our customers regarding warranty contracts.

Fueling Innovation with Big Data

Even with all we have accomplished, we've just begun to tap big data to fuel innovation, and we're already victims of our own success. By that we mean other groups within the company see what big data is doing for engineering and they're clamoring for access to this powerful resource. We're excited about the opportunities big data offers and the value of Hadoop in pursuing those opportunities. And as with big data, we have just begun to tap the robust functionality and enterprise-level scalability of Control-M in powering our journey of innovation.

Reduce costs and improve service delivery with faster Hadoop implementations with Control-M for Hadoop.