HighByte

MANUFACTURING TECHNOLOGY INSIGHTS

Vendor Viewpoint

Industrial DataOps: The Missing Link for Industry 4.0

Author: Tony Paine, HighByte Co-Founder & Chief Executive Officer

This article first appeared in its original form on manufacturingtechnologyinsights.com in 2020.

Summary
I've been developing solutions for the industrial automation industry for the past 25 years. In the last few years, I've witnessed the industry experience a dramatic step-change due to the rising adoption of Industry 4.0. However, I've also witnessed companies continue to struggle with the move from pilot to production due to gaps in their data infrastructure. As founder and CEO of HighByte, I'm on a mission to provide the critical infrastructure that bridges the gap between the collection of raw industrial data and the many disparate applications that need to make use of refined industrial information. I believe Industrial DataOps is the missing link for achieving the intended scale, speed, and impact of Industry 4.0 for modern manufacturers.

Current situation
Industry 4.0 has introduced new types of applications that solve a variety of important use cases, including enterprise-wide visualization, asset management, and supply chain optimization—just to name a few.

Many manufacturers we speak with are currently focused on deploying advanced analytics for predictive maintenance. By combining existing automation with new cost-effective IoT sensors, specific machine data can be used to predict the right time to perform maintenance, enabling manufacturers to reduce the cost of machine maintenance and eliminate unplanned downtime.

This same data—perhaps correlated and interpreted a little differently—can also be used for other purposes. For example, a manufacturer may want to analyze machine data to improve productivity and optimize the use of one or more assets across multiple sites. Outside the plant floor, analytics built into manufactured products can provide customers with higher levels of satisfaction. Manufacturers can instrument the use of their products to better understand customer needs and ultimately provide a higher quality product and service. In short, analytics can provide a manufacturer with better insights into the entire business—from production to customer service—and allow the business to better plan for future demand.

You can see how these new insights could also be re-injected into other systems. So, in effect, we don't just want to collect data, model it, and provide value within a specific application, we also need to take these new learnings and continuously iterate to close the loop on how information is shared between various components in the architecture.

Challenges for innovation
Data preparation is challenging for any industry, but the unique, sophisticated qualities of industrial data make this process much more complicated for manufacturers. In fact, our research shows that data scientists are spending up to 80% of their time preparing data for analysis rather than actually analyzing it for business insights. But why?

First of all, the data collected for advanced analytics must be sourced not only from machines and products, but also from pre-existing systems, applications, and even offline technical documentation. These sources exacerbate the complexity involved with collecting, preparing, modeling, and validating data sets for correctness before being served to an analytic.

Also, applications and systems within a single site often come from multiple vendors deployed at different periods of time by different personnel. Raw data needs to be standardized, normalized, and assembled in real time before it can be used. Context-aware information is often added on an application-by-application basis, as each application has its own demands on the structure and format of the information it requires in order to provide value. Learning a new way of representing the same types of information is tedious and presents maintainability challenges as the environment changes over time.

And, of course, not every application should have open access to all data. Many manufacturers struggle with the ability to gain visibility and control over the flow of real-time data within their organization and to outside vendors.

The current approach

In consultation with the operations team and on an ad-hoc basis, data scientists are remodeling existing information for their unique analytics needs. They learn where the raw data lives and how to access it through standards like OPC UA or vendor-specific APIs. They learn how to normalize raw data by peeking into SCADA systems and their associated scripting technologies that prescribe how to massage the data. They learn about data context from these same applications, through technical documentation, or by talking to plant engineers and other subject matter experts. Then they develop their own custom code to pull all the types of data together and wrangle it into information that can then be served to an analytic. After a heavy lift, they have made progress. But now they need to rinse and repeat the cycle in order to create the next one.

Unfortunately, they have also created a maintainability problem for the organization with another set of custom code. They've built a brittle solution reliant on operations. They forever need to know any changes made in the operations layer and how these changes may need to be reflected in their work. The technical debt is magnified when we consider unknown future needs and applications.

Deploying a new app should be focused on the additional value it provides—not the repetitive efforts of configuring, segmenting, and massaging the same data for a different purpose. There must be a better way.

Solutions and suggestions

I believe that like successful data collection platforms at the bottom of the stack, we also need to abstract away how information is modeled at the top of the stack.

The market has recently introduced Industrial DataOps as a new abstraction layer in the technology stack that collects, models, and shares information between assets, products, processes, systems, and ultimately people. At its core, an Industrial DataOps solution is an application-agnostic modeling platform for cleansing, normalizing, standardizing and adding context to data before delivering a consistent and complete set of information to all consuming applications.

A modeling platform allows manufacturers to standardize a naming convention or unified namespace that is consistent across the enterprise—no more trying to understand how a piece of data maps into the entire solution. Manufacturers can ensure information is assembled from various data sources under a single set of conditions. This prevents the various consuming applications from performing this task themselves and not having the same exact same record of information due to timing or other anomalies.

An Industrial DataOps abstraction layer enables manufacturers to look to best of breed, purpose-built solutions that, when networked together, achieve the desired architecture. By decoupling their data infrastructure from their architecture, manufactures can avoid vendor lock-in. If a particular piece of the solution no longer meets their needs, they can replace just this piece and proceed without disruption.

Whether organizations want to take a bottoms-up approach starting at the sensor level or a top-down approach starting at the site, how users model data in an Industrial DataOps solutions is entirely flexible. They no longer need to own and maintain custom code as the environment evolves.

By creating standard models in an Industrial DataOps solution, manufacturers can transform cryptic data tag schemes, leverage access to existing operational data, add appropriate context, and make industrial information available to the rest of the enterprise. The result is a data architecture that can scale and adapt when and where changes are needed. This is the missing link to Industry 4.0.

About the Author

Tony Paine is the CEO of HighByte and is leading the design and development of the company's initial product offering.

Tony has had a passion for integrating software and hardware since his early childhood, when he developed an application that turned a rudimentary text editor into a word processor with generic print capabilities. He focused his education around this interest and earned a Bachelor of Science degree in Electrical Engineering, with a concentration in Computer Software and Hardware Design, from the University of Maine at Orono in 1996.

For the past 20 years, Tony immersed himself in industrial software development and strategy at Kepware, most recently serving as CEO. He led the company through a successful acquisition to PTC in 2016.

Tony has contributed to a variety of technical working groups, helping to shape the direction of standards used within the Automation Industry. As a strong advocate for STEM initiatives, he currently sits on the Dean's Advisory Council for the University of Maine's College of Engineering where he provides industry insight and evangelism around education in the area of technology.

About HighByte

HighByte is an industrial software development company in Portland, Maine building solutions that address the data architecture and security challenges created by Industry 4.0. We believe contextualized and standardized data is essential for Industry 4.0 to reach broad adoption. That's why we've launched HighByte Intelligence Hub—enabling manufacturers to securely connect, model, and flow valuable industrial data throughout their extended enterprise without writing or maintaining code. HighByte Intelligence Hub is the first DataOps solution purpose-built to meet the unique requirements of industrial assets, products, processes, and systems at the Edge. Learn more and request a free trial at https://highbyte.com.