ARUNDO

EDGE ANALYTICS AND THE INDUSTRIAL INTERNET OF THINGS

Houston, USA
1600 Smith St Suite 3730
Houston, TX 77002

Oslo, Norway
Wergelandsveien 7
0167 Oslo

Palo Alto
444 Ramona Street
Palo Alto, CA 94301

CONTENTS

OVERVIEW

Cloud-based analytics permeate heavy industries more and more each day. As the maturity of cloud infrastructure and cloud-based applications increases, however, owners and operators of industrial assets are realizing the constraints of the cloud.

Advanced analytics for remote assets and equipment suffer from physical, logistical, and business constraints when applying purely cloud-based solutions. These constraints include:

- the cost of constant communications connectivity for far-flung or mobile assets,
- the cost of cloud replication of local data storage, and
- the ability to drive timely, valuable business insigts from all of this data to inform actions and decisions at remote sites.

Increasingly, owners and operators of assets such as maritime vessels, oil platforms, remote pump or compressor stations, or other intermittently-connected assets, look for “edge” solutions that bring the power of advanced analytics to local environments. Even when they are primarily operating in a disconnected state. However, such edge solutions are not merely hardware or software. Useful edge computing applications must drive business value in a tangible way given the requirements and responsiveness of potential actions. The balance between using centralized, massive cloud computing resources vs. distributed computing and local actions is ultimately found in the time and insight required to take relevant action in the field.

THE LIMITS OF CLOUD COMPUTING FOR INDUSTRIAL ANALYTICS

For industrial users, the promise of cloud computing is to bring supercomputing capabilities - previously restricted to only the most important, complex, and critical analyses - to more mundane, widely distributed business applications.

This is most obvious in the increasing prevalence of cloud-based analytics of massive datasets:

- time series data from equipment and process historians,
- ERP and inventory data,
- safety or maintenance records
- vendor, partner, and OEM data
- GPS, weather, or other third-party data,
- decades of paper records slowly being digitized and indexed for future analytics projects, and
- other sensors, databases, and services too numerous to count.

This work is important. In major industries such as oil & gas and shipping, it can be transformative. Literally reshaping how companies operate significant components of their business. For many companies, just migrating data and analytics processes to the cloud will be the primary focus of digital transformation for the foreseeable future.

However, many companies in these industries face a significant structural disadvantage in bringing advanced analytics to some of their most critical assets. These are "edge" assets – those located in rugged, remote environments without a persistent internet connection. Real-time cloud data analytics makes less sense in the context of assets such as offshore oil platforms, remote pump or compressor stations, or ocean-going vessels.

However, the need for analytics in such assets is critical. In fact, these assets are disproportionately capital-intensive. Often, the very fact of their remote environment requires a level of engineering that more accessible assets do not need. This creates even more value-added opportunities for analytics.

For example, an average offshore oil platform creates 1-2 terabytes of data every day (a number which is estimated to double every three years). Pushing this data to the cloud via satellite at typical speeds would take 12 days just for 1 day's worth of data.[1] It is simply not feasible to perform time-dependent analytics across massive datasets with such constraints. Yet these oil platforms represent a significant asset with enormous value potential from improved performance and uptime. Consequently, analysts such as Gartner predict that in less than three years, more than 75% of data consumed by industrial analytics applications will never leave the edge.[2]

Edge computing represents a distributed architecture whereby data is accessed, aggregated, filtered and analyzed at the periphery of the network. Edge technology augments both cloud and local control systems in a cost-effective manner.

1 Hand, Aaron. "Oil and Gas at the Edge," Automation World, 3 September, 2017. https://www.automationworld.com/oil-and-gas-edge

2 Van der Meulen, Rob. "What Edge Computing Means for Infrastructure and Operations Leaders," Smarter With Gartner, 3 October, 2018. https://www.gartner.com/smarterwithgartner/what-edge-computing-means-for-infrastructure-and-operations-leaders/

THREE LEVELS OF EDGE ANALYTICS

As discussions of edge analytics become more prevalent in conferences, papers, and meetings about the industrial internet of things (IIoT), there is some confusion about what this term means. There are three basic flavors of edge analytics:

1. Intelligently ingesting data from edge-based equipment and machines, even without persistent internet connectivity, and sending data to cloud-based analytical software whenever an internet connection is available
2. Performing rules-based analytics (formula-based calculations) locally
3. Performing advanced analytics (compute-intensive operations such as machine learning algorithms or advanced simulations) locally

These various levels of edge analytics require different investments in hardware and software at the edge and in the cloud. Depending on the end use case, asset owners and operators must choose the best investment level for their specific situation.

INTELLIGENTLY INGESTING EDGE DATA

As discussed above, regularly backhauling 100% of data generated by local instrumentation is cost-prohibitive, time-intensive, and unlikely to significantly improve decisions and actions for local operators.

Intelligent edge ingest focuses on pulling only the most critical data. For instance, the sensors related to the most "interesting" remote pump stations – and streaming this data based on the availability of communications (even if this connectivity may be irregular). The data then goes to a cloud-based repository where cloud analytics can be performed, with results pushed back to the edge or to centralized reporting workflows. The key technical challenges in this process include accessing and ingesting sensor values regardless of local architecture, buffering data, and checking for communications availability. Arguably, the most critical part of this challenge is to enable maximally equipment agnostic sensor data ingestion.

While long-standing efforts to standardize equipment control system protocols (e.g., OPC-UA) continue, there will remain significant fragmentation in the near-term. This is partially due to temporary inertia among equipment manufacturers to open up data access at reasonable commercial terms. This inertia will give way to increasing pressures from asset operators (who demand integrated visibility across all types of assets regardless of origins) as well as a growing open source community for industrial connectors. The computing hardware requirements are relatively cheap and minimal, with the largest expense typically related to specialty enclosures for extreme locations.

PERFORMING RULES-BASED ANALYTICS AT THE EDGE

Simple formula-based calculations – arithmetic, algebraic, or even calculus-based – can be performed before sensor values leave an edge location.

Compute rules may be installed or created on edge hardware so data is locally analyzed as it is acquired (e.g., a formula for efficiency may take input, output, and energy sensor values and return a single efficiency value continuously). The outputs of such calculations may then be streamed to the cloud, similar to raw sensor values, or used locally depending on the use case. The computing hardware requirements are still relatively cheap and minimal, aside from any specialty enclosures.

PERFORMING ADVANCED ANALYTICS AT THE EDGE

The most powerful use of local sensor data often requires analytics beyond simple rules-based calculations. Deploying machine learning models or compute-intensive simulation analytics to the edge (for instance, to classify or detect abnormal equipment behavior in real-time) can be challenging, however.

An integrated solution that can ingest data from local sensors and deploy, store, and run cloud-based analytical models at the edge is often the best answer. Of course, depending on the complexity and compute-intensity of advanced models at the edge, the local computing hardware requirements may be significantly different than the relatively lightweight requirements for streaming data or performing more simple calculations.

GETTING STARTED WITH EDGE ANALYTICS

Navigating the technology decisions to implement edge analytics may span different departments in a typical organization.

For instance, a control and instrumentation group may govern field data access, an IT organization may drive global cloud decisions, while an analytics or digital group may want to implement analytics applications at the edge. As with many decisions that cross functions or organizations, fast, low-risk approaches are often the best way to get started.

The piece parts of an edge computing system can be rolled out separately:

1. Identify and educate the most relevant stakeholders as to what you intend to try.
2. Identify a set of sensors or programmable logic controller (PLC) tags that could be helpful for a field operator or decision-maker to monitor or analyze on a continuous basis. This could be an equipment type, such as remote pumps, or an asset type, such as an ocean-going vessel.
3. Install the hardware and software locally to connect to these sensors or tags. Typically this should cost less than $2,000, and be simple enough to install by anyone who knows how to use a computer.

This should avoid major budgetary or procurement cycles.

4. Stream the data to an existing cloud tenant, or set up a free trial with any of the major cloud providers.
5. Now that data is streaming from edge to cloud, showcase this success for the relevant stakeholders. Determine how this ability could improve your existing operations processes.
6. Finally, drive to the more advanced analytical goals, potentially with simple local calculations as an intermediate step. This may involve more protracted work with analytics teams or third-party providers but should be done in the context of fast implementation and immediate business value.

CONCLUDING THOUGHTS

Edge computing and edge analytics are poised to change an already complex and evolving IIoT landscape for heavy industrial operations. There are simple, easy ways for companies to experiment with edge computing before making significant investments. Such "no-risk" moves represent one of the best ways to drive corporate analytics programs to focus on business value in general, as edge analytics only make sense in this context.

AUTHOR

Amitav Misra is General Manager Americas and VP Marketing at Arundo Analytics. He works closely with industrial equipment manufacturers and major asset owners seeking equipment monitoring solutions. He previously worked at EnerNOC Inc. and McKinsey & Co. Contact him at amitav.misra@arundo.com

ABOUT ARUNDO

With offices in Oslo, Houston and Silicon Valley, Arundo Analytics provides cloud-based and edge-enabled software for the deployment and management of enterprise-scale industrial data science solutions. Arundo's software allows industrial companies and other organizations to increase revenue, reduce costs and mitigate risks through machine learning and other analytical solutions that connect industrial data to advanced models and connect model insights to business decisions.

In 2016, Arundo graduated from Stanford University's StartX accelerator program, and subsequently received investment from the Stanford-StartX Fund. In 2017, Arundo was named to the MIT STEX25 by the Massachusetts Institute of Technology Startup Exchange (MIT STEX). MIT STEX25 recognizes select companies from a pool of more than 1,000 MIT-connected startups as being particularly well-suited for industry collaboration based on technical and commercial success.

For more information, please visit www.arundo.com

If you would like to explore opportunities to connect, analyze, and act upon your remote field data, please contact us.

Americas
Amitav Misra
amitav.misra@arundo.com

EMEA
Edoardo Jacucci
edoardo.jacucci@arundo.com