Intelligence at the Edge Part 1: The Edge Node

Ian Beavers
Analog Devices, Inc.

The industrial Internet of Things (IoT) encompasses the broad transformation underway that will make pervasive sensing across connected machines not just a competitive advantage, but an essential fundamental service. The industrial IoT starts with the edge node, which is the sensing and measurement entry point of interest. This is where the physical world interacts with computational data analytics. Connected industrial machines can sense a wide array of information that will be used to make key decisions. This edge sensor is likely far removed from the cloud server that stores historical analysis. It must connect through a gateway that aggregates edge data into the internet. Ideally, the edge sensor node is unobtrusive within a small nominal form factor to easily deploy in space constrained environments.

Sense, Measure, Interpret, Connect

In this first of a multipart industrial IoT series, we will break down and explore the fundamental aspects of the edge node sense and measurement capabilities within the larger IoT framework: sensing, measuring, interpreting, and connecting data, with additional consideration for power management and security. Each portion presents a unique set of challenges. Smart partitioning of the edge node can be key to a successful implementation. In some cases, ultralow power (ULP) is the most important performance metric. The vast majority of potential data may be filtered when the sensor wakes from sleep mode during key events.

Sensors form the front-end edge of the industrial IoT electronics ecosystem. Measurements transform the sensed information into something meaningful such as a quantifiable value of pressure, displacement, or rotation. The interpretation stage is where edge analytics and processing transforms the measured data into an actionable event.[1] Only the most valuable information should be connected beyond the node into the cloud for predictive or historical processing. All along the signal chain, the data can be rejected or filtered based on initial limits of acceptability. Ideally, the sensor node should only send information that is absolutely necessary and should make critical decisions as soon as key data is available.

The edge node must be connected to the outside network, either through a wired or wireless sensor node (WSN). Data integrity remains key in this block of the signal chain. Optimum sensed and measured data is of little value if the communication is inconsistent, lost, or corrupted. Missing data via communication cannot be an option. Electrically noisy industrial environments can be harsh and unforgiving, especially for radio frequency communication in the presence of high metal content. Therefore, a robust communication protocol must be designed as a forethought during system architecture design.

Power management for ULP systems starts with regulator component selection for maximum efficiency. But, as edge nodes may also wake and sleep with a rapid duty cycle, the power-up and power-down time should also not be ignored. An external trigger or wake-up command aids in the ability to quickly alert the edge node to begin sensing and measuring data.

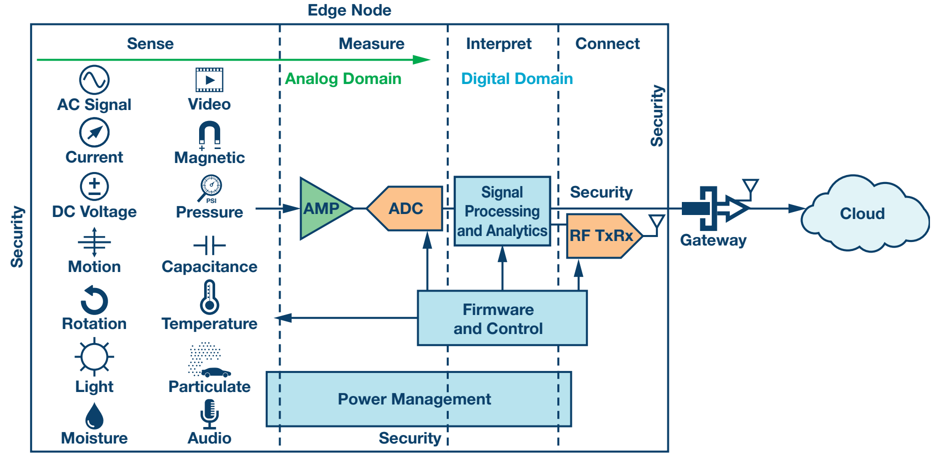

Figure 1. An edge node device provides the intelligence to sense, measure, interpret, and connect to an internet gateway to the cloud. The data can be preprocessed with some form of analytics before it is transmitted for deeper data mining intelligence.

Data security must also be a consideration for an industrial IoT system. Not only does the data protection within the edge need to be secure, but its access to the network gateway must also be protected from malicious intent. An edge node must not be allowed to be spoofed to gain network access for nefarious activity.

Intelligence Starts at the Edge

There are a legion of sensing solutions at the edge, which may not just be a single discrete device. The edge may be a plurality of various concurrent unrelated data acquisitions. Temperature, sound, vibration, pressure, humidity, motion, pollutants, audio, and video are just some of the variables that can be sensed, processed, and sent to the cloud through a gateway for further historic and predictive analysis.

It is not a hyperbole to say that sensors are the backbone of industrial IoT.[2] But it might be more accurate to say they're the central nervous system for extracting insights. The edge node sense and measurement technology is the birthplace for the data of interest. If bad or incorrect data is faithfully recorded at this stage in the solution chain, no amount of postprocessing in the cloud can reclaim the lost value.

Mission critical systems, such as healthcare and factory line-down monitoring with high stakes outcomes, require robust integrity of quality data measurements. Data quality is paramount. False positives or omissions can be costly, time consuming, and potentially life threatening. Costly errors eventually result in unplanned maintenance, inefficient labor use, or having to disable the IoT system entirely. Intelligence starts at the edge node where avoidance of the old adage still applies—garbage in, garbage out.

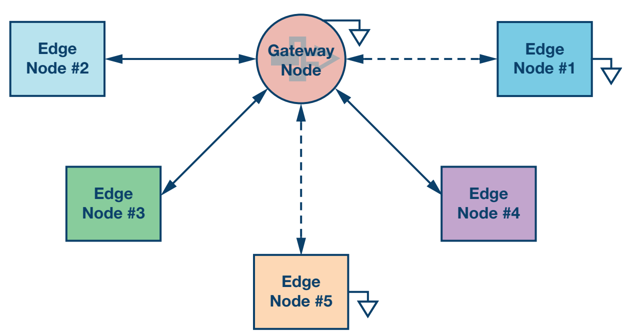

Figure 2. Many edge node outputs, both wired and wireless, can autonomously connect to a gateway node to be aggregated before transmission to a cloud server.

With Access to Troves of Data, Comes Great Responsibility

In legacy signal chain solutions without edge node intelligence, data stays data. An unintelligent node never helps generate wisdom and knowledge to make actionable decisions.[1] There can be large quantities of raw, low quality data that have no impact on the system performance of interest.[3] It can be power hungry and bandwidth intensive to convert and send all of this data to an eventual cloud storage destination.

By contrast, intelligent smart partitioning edge node sensing and measuring turns data into actionable information. An intelligent node lowers overall power consumption, lowers latency, and reduces bandwidth waste.[4] This enables a move from a reactive IoT with a long latency to both a real-time and predictive IoT model. Basic analog signal chain circuit design philosophy still applies to the IoT. For complex systems, deep application expertise is often needed to interpret the processed data.

Optimized Smart Partitioning Maximizes Cloud Value

Only the most important measured information needs to be sent through the gateway to the cloud for final processing. In some cases, a majority of the data is completely unimportant.[5] However, system data that is time critical with a need for a local real-time decision should be acted upon long before it is aggregated at a distant point with remote access. In contrast, information that leverages historical value with prediction models to influence long-term insights makes an ideal application for cloud processing. Archiving data into mass databases for retroactive processing and decisions plays to the strengths of powerful cloud processing and storage.[6]

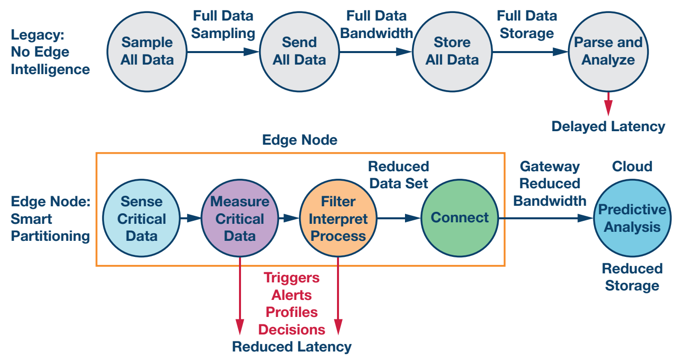

Figure 3. Smart partitioning at the edge node solves new challenges that could not otherwise be solved before. Leaner processing and intelligence further ahead in the signal chain allow for a more efficient total IoT solution.

Living on the Edge with Real-Time Decisions

IoT sensors are predominantly analog. The specific industrial application requirements will dictate the dynamic range and bandwidth of the sensor that will be required at the front end of the edge node. The front end of the signal chain will be within the analog domain before the signal is converted to a digital representation and transmitted outside the edge. Each component in the analog signal chain has the potential to limit the overall performance of the edge node if not selected properly. The dynamic range will be the delta between the full-scale sensor of interest relative to the noise floor or next highest undesired signal.

Since IoT sensors are typically looking for both known and unknown activities, an analog filter does not always make sense. Digital filtering is performed after the signal is sampled. Unless an analog filter is used at the front end of the sensor, harmonics of the fundamental or other spurious signals can fold into the sensed information and compete in power with the signal of interest. Therefore, planning for unanticipated sensed signals in both the time and frequency domain during the design phase will prevent unwanted artifacts from showing up in the measured data.

The sensed information is typically measured with an ADC that is next in the signal chain. If the IoT edge node is designed using discrete components, care should be taken to choose a measurement ADC that does not reduce the dynamic range of the sensor. An embedded ADC's input full-scale range is usually matched well to the sensor output amplitude. Ideally, the sensor output should consume nearly the entire ADC input range, within 1 dB, without saturating the ADC and getting clipped at the range limits. However, an amplifier stage may also be used to gain or attenuate the sensor output signal to maximize the ADC's own dynamic range. The ADC full-scale input, sample rate, resolution in bits, input bandwidth, and noise density will all contribute to the signal measurement performance of the edge node.

A front-end amplifier can either be embedded in the measurement of the node or added as a discrete component ahead of the ADC. The gain, bandwidth, and noise of the amplifier can also enhance the performance of the edge node.

The measurement ADC after the sensor in the signal chain is often one of two sampling architecture types: Nyquist rate or a continuous time Σ-Δ (CTSD), with the latter being more prevalent with embedded ADCs. A Nyquist rate ADC will have a nominal flat noise floor equal to half the sample rate frequency, or $f_s/2$. A CTSD uses an oversampling rate with a notched pass-band that pushes the noise outside the bandwidth of interest for increased dynamic range. The measurement ADC architecture and its resolution are key to understanding the analog bandwidth and dynamic range of the edge node.

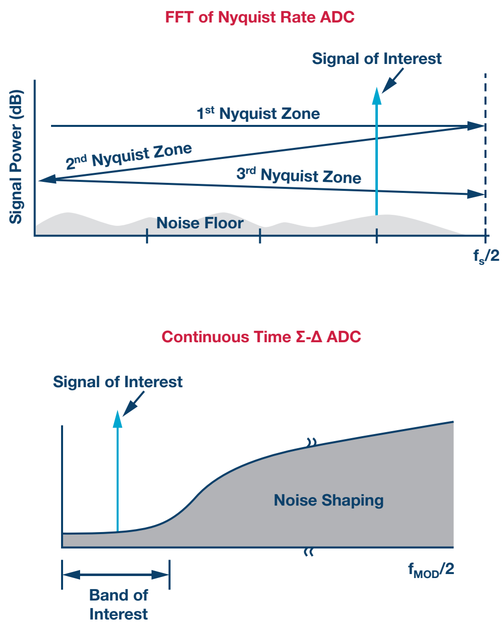

Figure 4. Without a front-end analog filter on an IoT sensor, a Nyquist rate ADC will fold higher-order frequencies beyond the 1st Nyquist zone back into the bandwidth of interest. By contrast, a CTSD ADC architecture with an oversampling modulation clock uses noise shaping to allow high dynamic range within a band of interest. The CTSD is less sensitive to signal aliasing as it provides inherent filtering.

For example, in the frequency domain, the noise density per unit bandwidth of 1 Hz will be based upon the SNR of the ADC and how wide the noise is spread across the sampled spectrum of the ADC. In a Nyquist rate ADC, the noise spectral density (per 1 Hz bandwidth) = 0 dB – ADC signal-to-noise ratio (SNR) – 10 × log($f_s/2$) where $f_s/2$ is the sample rate divided by two or a single Nyquist zone of the ADC. The ideal SNR can be calculated as SNR = 6.02 × N + 1.76 dB, where N is the number of ADC bits. However, the actual SNR of an ADC includes the nonidealities of transistor and semiconductor processing, including electrical noise and transistor level component imperfections. These nonlinearities will degrade the SNR performance below the ideal, so check the ADC data sheet for the SNR performance of interest.

The dynamic range of the edge node will be composed of the dynamic range of the sensor, the amplification of the signal, if needed, and the ADC full-scale dynamic range. If the full-scale sensor output signal does not reach within 1 dB of the ADC full-scale range input, then some portion of the ADC dynamic range will be left unused. Conversely, an overranged ADC input from the sensor will distort the sampled signal. Amplifier bandwidth, gain, and noise will also be part of the consideration for the dynamic range of the edge node. The electrical noise of the sensor, amplifier, and the ADC combined will be the square root of the square sum of each rms component.[7]

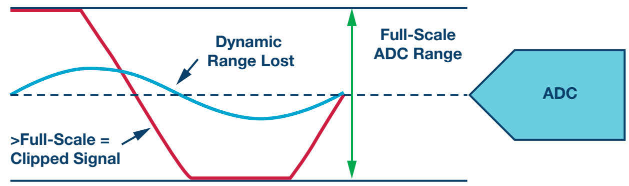

Figure 5. An example of a sensor signal output amplitude that does not match the input full scale of the ADC and dynamic range is lost (blue). An amplifier is needed to maximize the dynamic range of the sensor while preventing ADC saturation (red). Signal matching must consider the bandwidth, dynamic range, and noise of the entire edge node signal chain.

Smart Factory

One application that will be important within the industrial IoT is machine vibration condition monitoring. New or legacy machine equipment can have key mechanical components, such as rotating shafts or gears, mounted with high dynamic range MEMS accelerometers.[8] These multiaxis sensors sample the vibratory displacement of machinery in real time. Vibration signatures can be measured, processed, and compared to an ideal machine profile.[9] In a factory, the analysis of this information aides increased efficiency, reduces line-down situations, and can predict mechanical failures in advance. In extreme cases, a machine with a rapidly deteriorating mechanical component, that would otherwise induce further damage, can be immediately shut down.

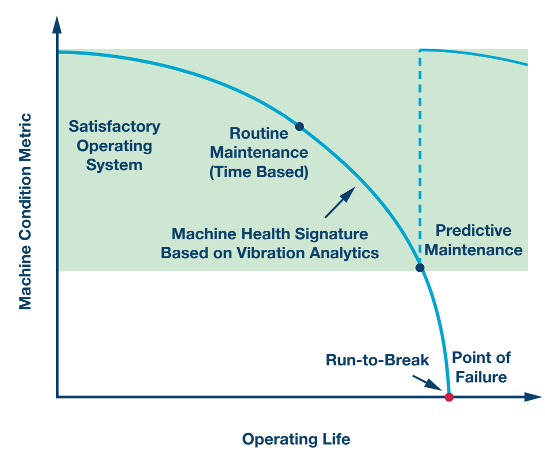

Figure 6. Although routine machine maintenance can be performed at regular time intervals, it is often not done with intelligence about the condition of the machine.[10] By analyzing the vibration performance of specific machine operations, a predictive point of failure and maintenance milestone can be alerted at the edge node.

Decision time latency can be drastically reduced by enabling edge node analytics. An example of this can be seen in Figure 7, where a MEMS sensor warning threshold limit is exceeded and an alert is immediately sent. If the event is extreme enough to be deemed critical, the node may be given authority to automatically disable the offending equipment to prevent a time sensitive catastrophic mechanical breakdown.

Alternately, a trigger signal may be invoked to enable another sense and measurement node, such as one on a secondary machine component, to begin interpreting data based on the 1st event. This reduces the total data set of sampled data from the edge nodes. In order to determine any vibration abnormality from nominal, the front-end node must be designed with the required performance for detection. The dynamic range, sample rate, and input bandwidth of the sense and measure circuit should be more than sufficient to identify any excursion event.

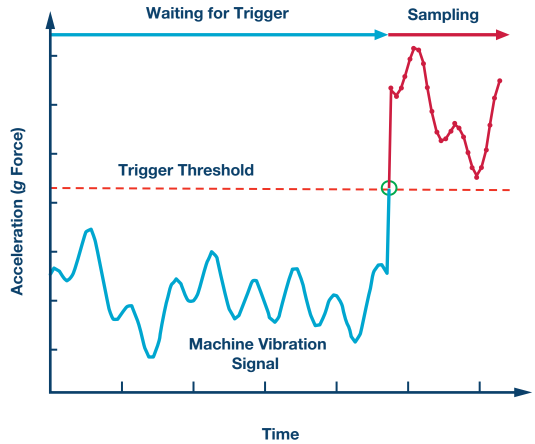

Figure 7. A time domain representation of sampled machine vibration data where a comparator threshold can determine whether or not the sensed and measured data is communicated beyond the edge. A lower power state can be maintained to filter a majority of the information until data preponderance is achieved by a threshold crossing event.

Smart City

A different industrial IoT edge node application is that of a smart city industrial camera with embedded video analytics. Smart city defines the urban mission to integrate a myriad of information and communication points into a cohesive system to enable management of a city's assets. A common application is to provide parking space vacancy alerts and occupancy detection. At the time of commissioning, each camera has a predetermined field of view. Boundary edge detection can be defined and used within the analytics to identify a variety of objects and their motion. Not only can historical object movement be analyzed, but due to object trajectory, a predicted path can also be computed at the edge using digital signal processing (DSP) algorithms.

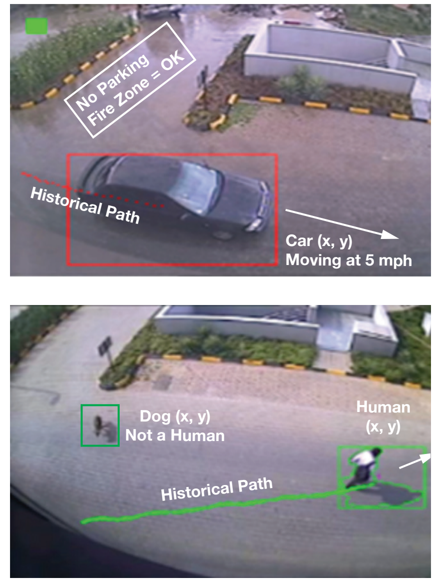

Figure 8. Using edge node video analytics, object type detection, trajectory, and boundary crossing can be determined in a low power system without sending full bandwidth video data to the cloud for analysis. Only a timestamp with breadcrumb object coordinates and type need to be communicated.

In a similar vein to frequency filtering, the full bandwidth of a video analytics frame is typically not needed for end processing. Often, when not used for security purposes, only a small subset of the complete video frame is required. Most of the visual data from frame to frame is static on a fixed mounted camera. The static data can be filtered. In some cases, only a count of boundary crossings or the movement coordinates of the object of interest needs to be analyzed. The reduced subset can be communicated as a breadcrumb coordinate to the next gateway in the signal chain.

Edge node video analytics can provide many filtered interpretations to differentiate object types—car, truck, bicycle, human, animal, etc. This decimation reduces data bandwidth and computational power that otherwise would be needed within a cloud server to analyze the full frame rate video data sent downstream.

Indoor camera applications may identify the number of people that cross an entry boundary and adjust the lighting, heating, or cooling for a room. In order to be visually effective in extreme lighting conditions or other challenging illuminations such as rain, a high dynamic range camera may need to be used in outdoor cameras. A typical 8-bit or 10-bit per pixel imaging sensor may not provide sufficient luminance dynamic range that is independent of illumination across all detection scenarios. In contrast to viewing fast motion sports at a 240 Hz refresh rate, a slower frame rate can be used to monitor the activity on an industrial analytics camera.

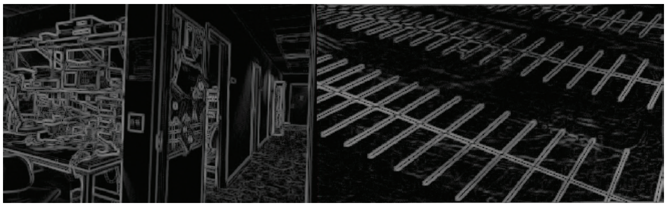

Figure 9. High dynamic range imagers with DSP object detection algorithms at the edge node can determine movement and boundary intrusions, even in low lighting conditions. This example uses visual contrast to define edge detection for an indoor factory/office (left) and an outdoor parking lot (right).

Platform Level Solutions

The ADT7420 is a 4 mm × 4 mm digital temperature sensor with breakthrough performance that contains an internal 16-bit ADC with resolution to 0.0078°C, consuming a mere 210 µA. The ADXL362 is an ultralow power 3-axis MEMS accelerometer that consumes only 2 µA at a 100 Hz sample rate in motion triggered wake-up mode. It does not use power duty cycling, but rather employs a full bandwidth architecture at all data rates, which prevents aliasing of input signals. The ADIS16229 is a dual-axis, 18 *g* digital MEMS vibration sensor with embedded RF transceiver. It also provides on-board frequency domain signal processing with a 512 point digital FFT capability.

A DSP enabled Blackfin low power imaging platform (BLIP)[11] allows rapid prototyping for industrial vision designs based upon proven digital signal processing tools. A library of optimized software deliverables allows equipment manufacturers an out-of-the-box solution for motion sensing, people counting, and vehicle detection.

References:

[1] Colm Prendergast. "Smart Partitioning and Value Creation in a Connected World." Keynote IoT Address: IESA Vision Summit 2015.

[2] Stephen Lawson. "IoT Keeps Pushing Analytics Closer to the Edge." *PCWorld:* IDG News Service, 2016.

[3] Lisa Morgan. "Edge Analytics an Antidote to IoT Data Deluge." *InformationWeek:* UBM Electronics, 2016.

[4] Daniel Kirsch. "The Value of Bringing Analytics to the Edge." Hurwitz & Associates Services, 2015.

[5] Jason Stamper. Why IoT Is Driving Analytics to the Edge of the Network. 451 Research, 2015.

[6] Steve Nelson. "Exploring the Internet of Things from End to End." Element14: Newark Electronics, 2014.

[7] Umesh Jayamohan. "Understanding How Amplifier Noise Contributes to Total Noise in ADC Signal Chains." *Analog Dialogue,* February 2013.

[8] Robert Randall. *Vibration-Based Condition Monitoring*. New South Wales, Australia: University of New South Wales, 2010.

[9] Ed Spence. "Bring the Benefits of MEMS Accelerometers to Condition Monitoring." *Electronic Design*, Penton Publishing, 2016.

[10] Jamie Smith. "Smart Edge Devices for the Industrial Internet of Things." ARC Industry Forum, 2015

[11] Blackfin Low Power Imaging Platform (BLIP). Analog Devices, Inc., 2014.

About the Author

Ian Beavers is a product engineering manager for the Automation Energy and Sensors team located at Analog Devices, Greensboro, NC. He has worked for the company since 1999. Ian has over 19 years of experience in the semiconductor industry. Ian earned a bachelor's degree in electrical engineering from North Carolina State University and an M.B.A. from the University of North Carolina at Greensboro. He can be reached at *ian.Beavers@analog.com*.

Analog Devices, Inc.
Worldwide Headquarters

Analog Devices, Inc.
One Technology Way
P.O. Box 9106
Norwood, MA 02062-9106
U.S.A.
Tel: 781.329.4700
(800.262.5643, U.S.A. only)
Fax: 781.461.3113

Analog Devices, Inc.
Europe Headquarters

Analog Devices GmbH
Otl-Aicher-Str. 60-64
80807 München
Germany
Tel: 49.89.76903.0
Fax: 49.89.76903.157

Analog Devices, Inc.
Japan Headquarters

Analog Devices, KK
New Pier Takeshiba
South Tower Building
1-16-1 Kaigan, Minato-ku,
Tokyo, 105-6891
Japan
Tel: 813.5402.8200
Fax: 813.5402.1064

Analog Devices, Inc.
Asia Pacific Headquarters

Analog Devices
5F, Sandhill Plaza
2290 Zuchongzhi Road
Zhangjiang Hi-Tech Park
Pudong New District
Shanghai, China 201203
Tel: 86.21.2320.8000
Fax: 86.21.2320.8222

TA15307-0-8/17(A)

analog.com

ANALOG DEVICES

AHEAD OF WHAT'S POSSIBLE™