

Big Analog Data: The Big Data for Engineers and Scientists

A Three-Tier Strategy for Data Analysis from Sensor to Cloud

By Tom Bradicich, PhD, Fellow and Corporate Officer, National Instruments,
and Kamalina Srikant, B.Eng, Market Development Manager, Embedded Systems, National Instruments

Engineers and scientists worldwide are acquiring vast amounts of data at very high speeds. As shown in Figure 1, physics experiments can generate tens of terabytes in just a few seconds. Testing of jet engines or electric power turbines can generate similar amounts in a matter of hours. Even smart grid measurements can generate terabytes of data over the course of a month.

Capturing, analyzing, and sharing this data is a “Big Analog Data” problem. In other words, it combines conventional Big Data issues with the difficulties of managing analog data. To cope with these challenges – and to harness the value in analog data sources – engineers need an end-to-end solution that can manage data from the sensor all the way to the cloud and beyond.

In this article we characterize Big Analog Data and explain how to harness this data using a three-tiered approach. We will then show how National Instruments (NI) is using Intel® technology to enable end-to-end solutions across these tiers. Finally, we will highlight applications in asset monitoring, automated testing, and scientific research.

Characterizing Big Analog Data

Connectivity in the emerging Internet of Things (IoT) is unleashing large amounts of analog data. As shown in Figure 2 (Page 10), this data is derived from a wide range of physical phenomena generated by people, nature, or machines. For example, a smart phone incorporates many analog phenomena such as sound, radio waves (cellular, Bluetooth, Wi-Fi, and GPS), vibration, light, touch, video, and orientation. When this smart phone is undergoing automated test at the end of its manufacturing line, a Big Analog Data problem exists.

In general, Big Analog Data is a subset of Big Data, which is commonly characterized by a combination of four “V’s” – Volume, Variety, Velocity, and Value. Another “V” that is becoming important is Visibility – globally dispersed enterprises need access to the data in multiple locations and across multiple levels of the organization. This is particularly true as growing data volumes prompt companies to distribute data to many more engineers and data scientists than in the past.

Big Analog Data is distinguished from other Big Data, such as that found in information technology (IT) systems, in two fundamental ways.

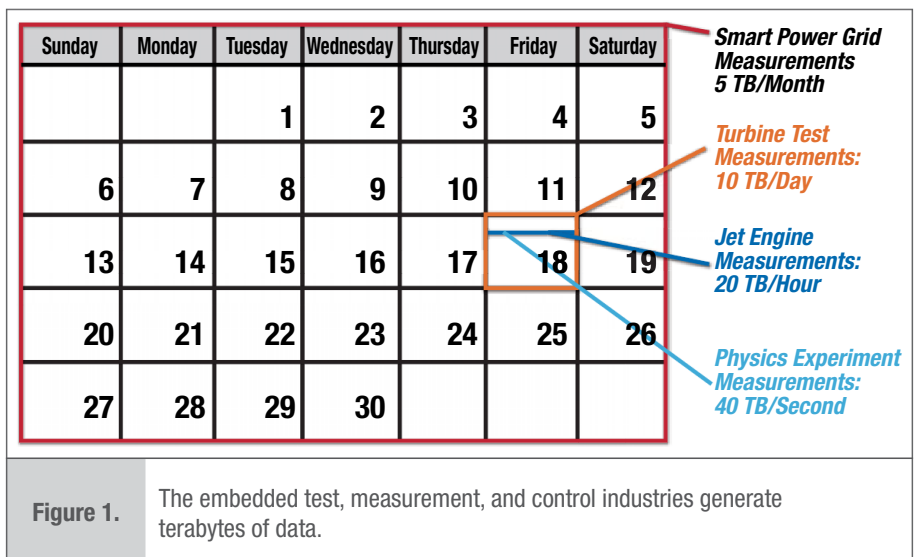


Figure 1.

The embedded test, measurement, and control industries generate terabytes of data.



First, it's faster since analog signals require digitizing at rates as fast as tens of gigahertz, often at large bit widths. Second, it's bigger because Big Analog Data information is constantly being generated from natural and man-made sources. Consider the unceasing electromagnetic waves all around us, offering information about the sunlight available for solar power (for example) or carrying data over wireless communications.

The Three-Tier Solution

In order to extract value from this vast quantity of complex data, organizations need advanced tools and techniques for data transfer, management, and analytics, as well as system management capabilities for the many data acquisition and automated test system nodes involved. These needs can be met with a three-tier architecture as depicted in Figure 3. These tiers come together to create a single, integrated solution adding insight from the real-time capture at the sensors to analytics in the back-end IT infrastructure.

Data starts in tier 1 at the sensor and is captured in tier 2 system nodes. These nodes perform initial analysis of the data in motion. Data is then filtered, and information deemed important flows across the edge to traditional IT equipment. In the IT Infrastructure tier, servers, storage, and networking equipment manage, organize, and further analyze the at-rest data. Finally, data is archived for later use.

Beneficial engineering, scientific, or business insight can be derived from the data throughout its life. For example, real-time analytics are needed to determine the immediate response of a motion control system. At the other end, at-rest data can be retrieved for analysis against newer in-motion data, for example, to gain insight into the seasonal behavior of a power-generating turbine.

National Instruments (NI) Ingredients

By working with an array of partners including Intel, NI is facilitating end-to-end solutions across all three tiers. Our partners supply a broad range of tier 1 sensors and actuators, spanning simple voltage and current transducers to cameras and accelerometers. These sensors directly connect to tier 2 system nodes, such as the NI CompactRIO and NI PXI systems.

NI CompactRIO is a rugged programmable automation controller for field-deployed embedded control and monitoring (Figure 4). NI CompactRIO systems consist of an Intel® Core™ processor for communication and general processing, a reconfigurable chassis with a user-programmable FPGA, and hot-swappable I/O modules. Their open architecture, low power, and small footprint make NI CompactRIO systems well suited to a range of industries like energy and manufacturing.

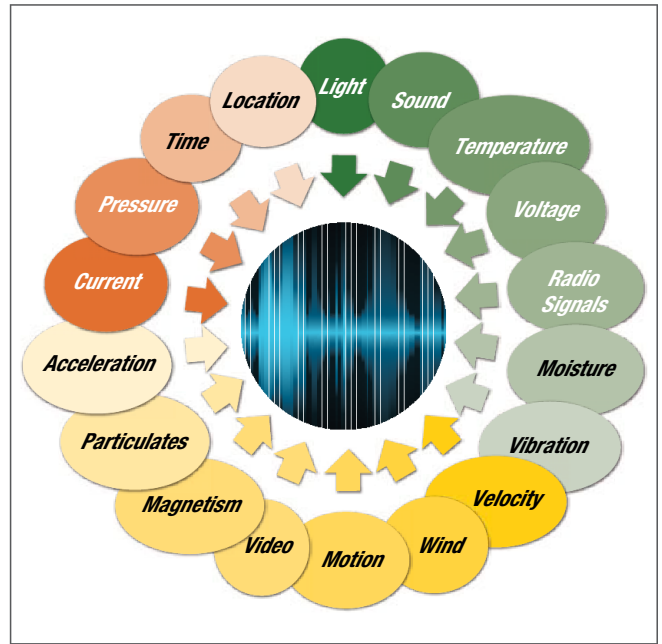


Figure 2. Big Analog Data comes from many sources.

NI PXI systems offer a rugged platform for test and measurement applications that can be driven either by a PC or by embedded control modules based on Intel® Atom™ and Intel Core processors (Figure 5). Based on the modular CompactPCI form factor, these systems add specialized synchronization buses, high-performance processing, and real-time measurement capabilities via embedded FPGAs. The NI PXI system is well established in manufacturing, military and aerospace, machine monitoring, automotive, and industrial test market segments, with more than 70 vendors providing I/O modules that support this interoperable specification.

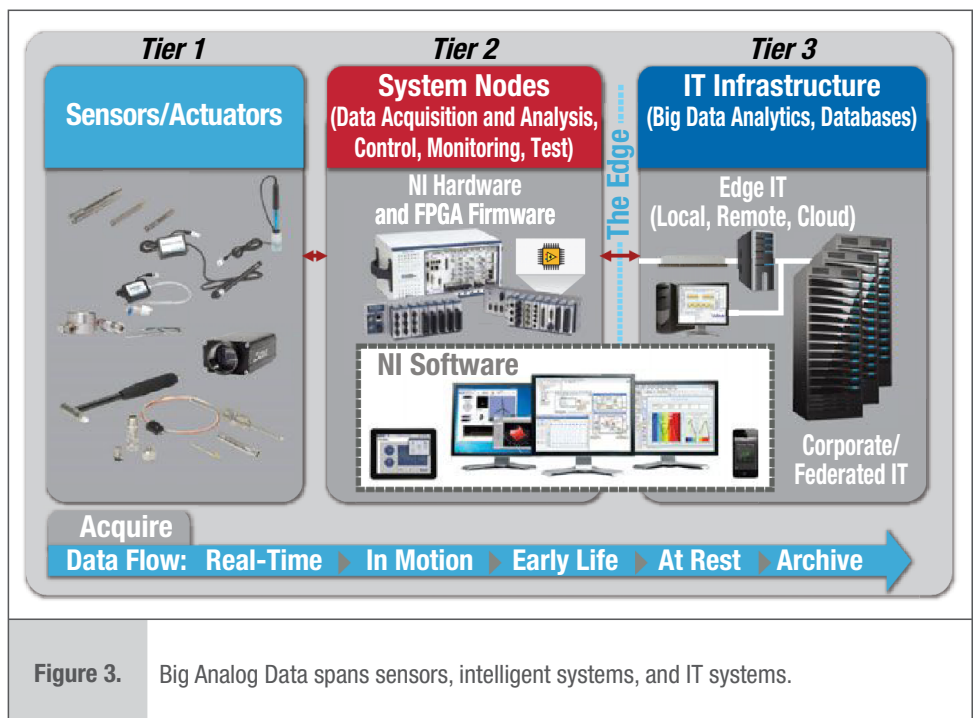


Figure 3. Big Analog Data spans sensors, intelligent systems, and IT systems.

Both hardware platforms are supported by a range of software tools, including the graphical NI LabVIEW software. This design tool abstracts the complexity of C- or VHDL-level programming for the FPGA within the target system. Using NI LabVIEW, developers can process the digitized signal in the real-time phase of the data flow, right at the point of data capture. This allows for the most immediate analytics, executed even prior to the data entering the Intel® processor in the intelligent system node.

NI LabVIEW and other tools also target the Intel processor-based controller (whether embedded or within a PC) to perform in-motion and early life analytics and visualization. The NI LabVIEW development and runtime environment allows users to implement all needed algorithms or choose from the vast number of NI software libraries that include general signal processing, image processing, and waveform signal analysis, as well as specialized custom algorithms and algorithms in third-party languages. NI DIAdem software is used to view synchronized mixed-signal data, historical trends, and compare multiple data sets. Along with NI DataFinder, this software is used for engineering data management, deep analytics, and robust presentation and visualization expressly designed for analog sources' data sets.

NI solutions also extend to tier 3 IT infrastructure. For example, the same NI DataFinder and NI DIAdem software that runs on the NI PXI intelligent nodes also runs on workstations and servers in tier 3 for scalable data and design analysis.

Also consider that Big Analog Data solutions typically involve many data acquisition channels connected to many system nodes. Hence, NI CompactRIO and NI PXI systems offer reliability, availability, serviceability, and manageability (RASM) features. This serviceability and management is similar to that needed for PCs and servers, and includes discovery, deployment, health status, updates, security, diagnostics, calibration, and event logging. These capabilities are critical for reducing the risks of integrating tier 2 solutions with IT infrastructure (See www.ni.com/rasm for more information about RASM in intelligent embedded and test systems).



Figure 4. National Instruments (NI) CompactRIO is a flexible platform for embedded control and monitoring.

Intel Ingredients

Intel components play a key role in powering three key functions for successful Big Analog Data solutions: data analytics, data management, and systems management. Specifically, Intel processors provide the performance needed for data analytics, software support for data management, and RASM for system management.

On the performance front, the Intel Atom and Intel Core processor families provide multi-core performance for sophisticated analysis of in-motion and early life data, as well as excellent energy efficiency for small form factors such as PXI blades. The superior performance of Intel processors also gives developers flexible design options. For example, developers can use a single Intel processor to analyze data, use the results to drive local system responses (such as motion control), and present local users with complex data visualization and rich user interfaces.

The baseline capabilities of these processors are enhanced by their integrated graphics and the availability of Intel® Virtualization Technology (Intel® VT). The advanced graphics engines built into these processors enable complex visualization at the tier 2 node. Intel VT – which is available in Intel Core processors as well as the new Intel® Atom™ processor E3800 product family – enables multiple operating systems (OSs) to run on the same hardware platform, allowing workload consolidation for efficient use and management of system resources. For example, Intel VT can be used to run a real-time OS for data analysis and motion control and Microsoft® Windows® for visualization and user interface, all on the same platform.

Intel VT also plays a critical role in tier 3 infrastructure, where it enables efficient server utilization. Supported across Intel® Xeon® processors, Intel VT allows IT staff to consolidate multiple servers on a single hardware unit and make efficient use of cloud resources.

Intel processors also give developers access to a vast software ecosystem, ranging from firmware to applications. This software compatibility greatly simplifies delivery of end-to-end Big Analog Data solutions. In addition to the NI DataFinder software mentioned earlier, Intel processors can run a wide range of software across tier 2 intelligent system

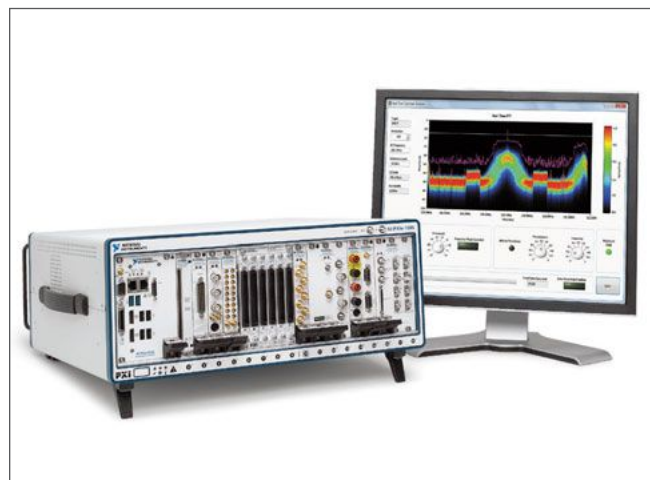


Figure 5. National Instruments (NI) PXI systems excel at test and measurement.



nodes and tier 3 infrastructure, greatly simplifying integration.

Finally, Intel enhances the entire end-to-end solution's RASM capabilities with a host of technologies, including Intel® vPro™ technology available in Intel Core and Xeon processors. Intel vPro technology encompasses a suite of security and manageability features, including:

- Intel® Active Management Technology (Intel® AMT), which enables remote monitoring, remediation, and repair
- Intel VT, which can securely separate critical tasks from less-crucial applications
- Intel® Trusted Execution Technology (Intel® TXT), which prevents unauthorized software from starting up

Intel and its subsidiaries provide a wide variety of other RASM solutions, both at the hardware and the software level. These include:

- Intel® Advanced Encryption Standard New Instructions (Intel® AES-NI), which accelerates encryption and decryption. Intel AES-NI is available on newer processors in the Intel Atom, Core, and Xeon processor families, including the new Intel Atom processor E3800 product family.
- Secure Boot technology, which ensures the system starts in a known good state. Versions of this technology are available on the 4th generation Intel® Core™ and Intel Atom processor E3800 product families.
- Error-correcting code (ECC) memory interfaces, which enhance data reliability. This feature is found on many Intel Core and Xeon processors, and is also available on the Intel Atom processor E3800 product family.
- Security and manageability solutions from Wind River and McAfee that ease administration and reduce complexity of the tier 3 IT infrastructure

Big Analog Data in Action

Ecosystems are critical to Big Analog Data solutions, and the NI Alliance Partner Network includes over 700 companies. Intimate relationships with ingredient providers and integrators across all three tiers of the solution allow NI to facilitate end-to-end solutions that reduce integration risk and time to value. Big Analog Data solutions incorporating NI and Intel ingredients are also available separately from providers like Avera Technologies, Bloomy Controls, Integrated Test and Management, Nexjen Systems, and VI Engineering.

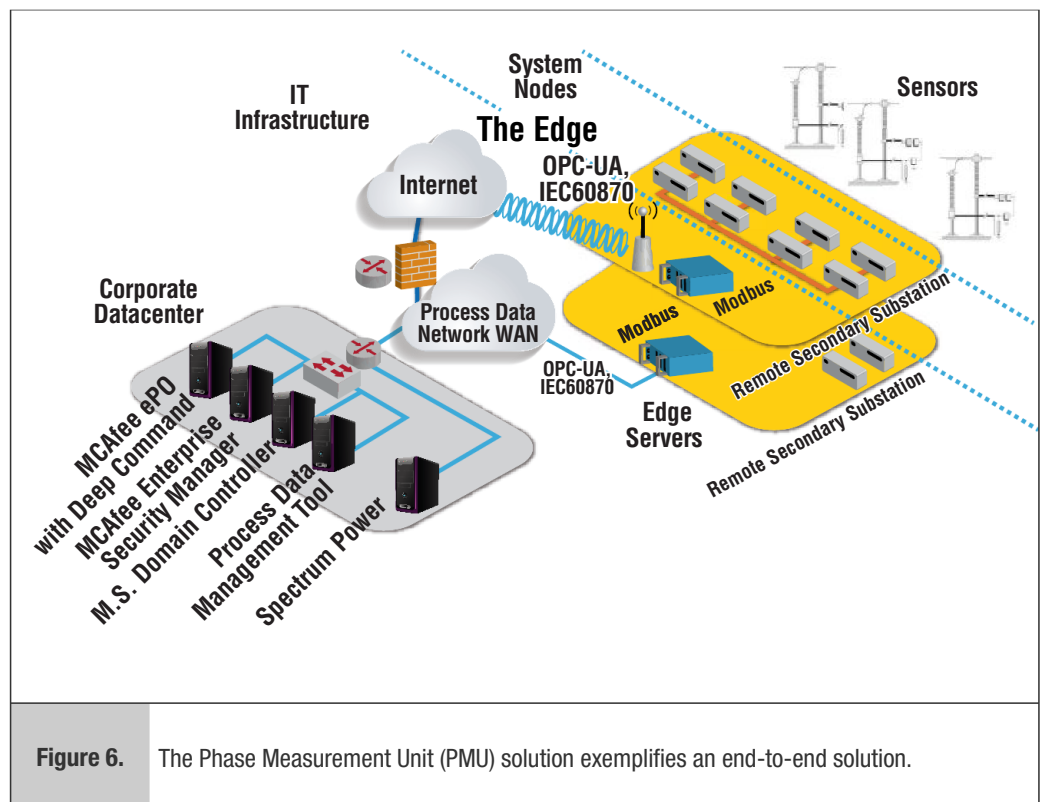


Figure 6. The Phase Measurement Unit (PMU) solution exemplifies an end-to-end solution.

As one example of an end-to-end solution, NI, Intel, Dell, and OSIsoft collaborated to offer the synchrophasor data management solution depicted in Figure 6. This solution combines high-performance phase measurement units (PMUs) from NI with Dell servers, storage, and networking equipment. The servers run phasor data concentrator (PDC) software from OSIsoft that collects and analyzes data from multiple PMUs. This solution enables advanced visualization, analytics, and early warning systems to help utilities detect evolving disturbances and avoid widespread blackouts (See ni.com/biganalogsdata for more details).

Unleashing Big Analog Data

The fastest and biggest Big Data – Big Analog Data – harbors great scientific, engineering, and business insight. To tap this vast resource, developers can turn to solutions powered by tools and platforms from NI and Intel. By collaborating closely with one another and a wide range of other partners, NI and Intel are enabling end-to-end solutions that span all three tiers of the Big Data hierarchy.



For more on National Instrument's solutions, see intel.com/SD-NI-products



For more on enabling Big Data and the Internet of Things with secure, manageable systems, visit intel.com/embedded-iot

Contact NI

National Instruments (intel.com/ea-ni) transforms the way engineers and scientists around the world design, prototype, and deploy systems for test and measurement, industrial control, energy distribution, and other applications. Using NI's design software and modular hardware, customers simplify development, increase productivity, and reduce time to market.