Solution Showcase

HPE and IoT Compute at the Edge

Date: March 2016 **Author:** Eugene Signorini, Senior Analyst

Introduction—IoT at the Intersection of IT and OT

In the Industrial Internet of Things, "things" stream large volumes of data from sources as diverse as wind farms, cars, copper mines, and hospitals. From a data perspective, the Internet of Things (IoT) can be viewed as gathering data from sensors and devices, deriving meaning from that data, and taking action on the data. Organizationally, IoT can be viewed as the intersection of information technology (IT) and operational technology (OT), from which the Industrial Internet of Things emerges.

The data from the "things," although once outside of IT's control, is now being merged with traditional IT data sources. To leverage insights from both IT and OT data, organizations must align compute power with the data's sources.

Sensor data can be collected, analyzed, stored (in the cloud or at a data center), and analyzed immediately or batch processed later. Alternately, streaming sensor data can be analyzed in real time, enabling companies to identify anomalies and patterns to immediately respond. The challenge is to understand what the data means, how it impacts your operations, and how long you have to act on the data. Different levels of urgency are required for monthly temperature averages versus things like turbines overheating. If immediate reaction is required, you can remove latency by moving data compute and analysis from the data center closer to the "things" at the edge, at the intersection of traditional IT and OT.

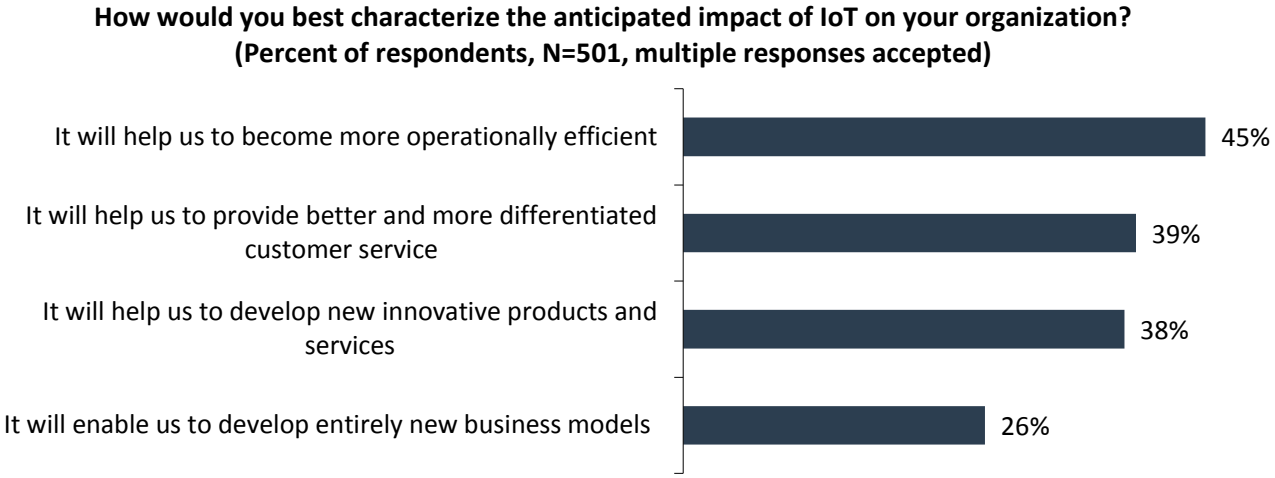
How Can Business Harness the Value from IoT?

ESG recently surveyed 633 IT and information security professionals representing enterprise-class (1,000 employees or more) organizations, which revealed that significant IoT activity is under way. While just 19% of firms claim they already have IoT initiatives underway, an additional 39% are currently developing IoT initiatives that will be launched in the next 24 months. This implies that organizations are in the process of understanding the immediate requirements and best practice approaches for IoT.

Why are so many companies moving ahead with IoT in light of the complexity that it represents? They see immediate business value in the form of operational efficiencies (45% of respondents), better and differentiated customer service (39%), creation of new products and services (38%), and development of new business models (26%).[1] However, in order to realize these benefits, organizations must build effective IoT solutions to empower the data from connected sensors and systems.

[1] Source: ESG Research Report, *2016 IT Spending Intentions Survey*, February 2016.

FIGURE 1. Anticipated Impact of Internet of Things



Source: Enterprise Strategy Group, 2016

Considerations for IoT Design: Information Technology and Operational Technology

As detailed in an earlier ESG white paper, *How to Choose an IT Platform to Empower Your Internet of Things*, the six key pillars of IoT are connectivity, compute, security, analytics, applications, and services. Following the flow of data from the endpoint to its eventual analysis provides a good perspective on the infrastructure requirements.

The source of the data is an endpoint (or sensor) that generates temperature, location, pressure, speeds, weight, color, or image data. The data may originate from a mobile connection in a vehicle, within a building from production or material handling equipment with RFID or barcode data, or a stationary manufacturing device connected via WiFi or wired. The type of data really depends upon the application that created it and may be converted or transformed. For example, temperature data may be converted into alarms for abnormal measurements. New data may be created as it moves along the path—for example, intermediate analysis may calculate averages, or it may be aggregated with data from other sensors to create data points. Therefore, the flow of data is not a straight line from the source to the destination, which differs from what we might expect in a traditional IT scenario, where network packets traverse from an endpoint (such as a personal computer) to the data center or to the Internet.

Instead, IoT design must address both information technology (IT) and operational technology (OT). The intersection of IT and OT has significant opportunities to drive new value to the organization. Together, IT and OT can transform operations such as manufacturing processes, coordination of networks of trains, or maintenance of a jet engine to get deeper insights into efficiency and reliability. Increasing it means that IT and OT become more interconnected.

This convergence presents organizations with challenges, as the characteristics of IT and OT differ across many dimensions, including business owners, primary function, technology lifecycle, reliability requirements, analytics needs, and the physical environment where the technology resides For compute, this requires a change in thinking about the What, Why, How, When, and Where.

IoT Compute Must Extend from the Edge to the Cloud

With IoT, compute units are found everywhere and not just in the data center. They may be found at endpoints (sensors with compute capabilities) or at the edge of the network where they process data before transmitting it to the data center, so a traditional viewpoint focused solely on the data center does not apply for IoT.

As mentioned, organizations should start rethinking their approach to compute within IoT initiatives by answering the What, Why, How, When, and Where:

- **What:** What is the business process that we are trying to impact? What type of data do we need, what can be collected, and what does not need to be saved?
- **Why**: Why do we need the data? For example, are we using the information to actuate a process in the physical world (i.e., the OT environment) or analyze data for insights (such as financial, customer, etc.)?
- **How:** How can we best apply computing resources (such as power and efficiency) to address the problems we are trying to solve?
- **When:** When do we need to collect, analyze, process, and act on the data?
- **Where:** Where is the best place to locate compute resources?

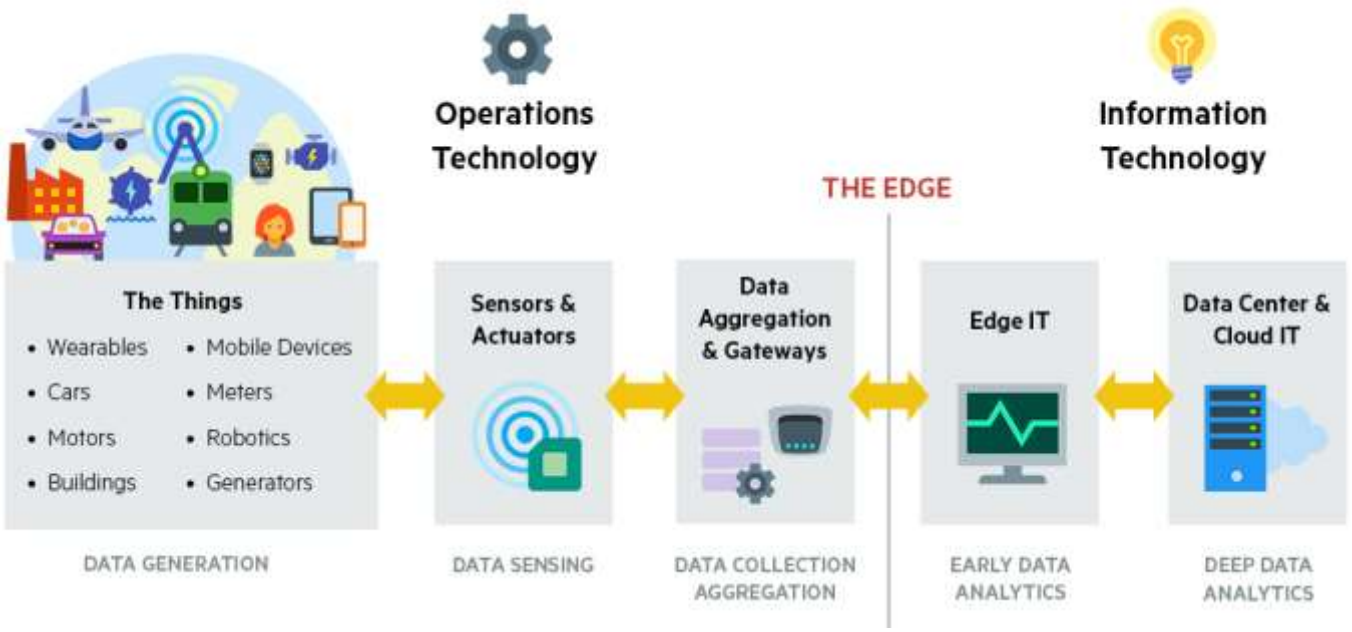
"Where" is listed as the last question, since the answers to the first four will likely determine where compute is executed: whether within a corporate data center, in the cloud, or at the edge. Although these are high-level questions, specific considerations can help organizations narrow down the answers. When answering these questions about compute capabilities within an IoT architecture, organizations should consider several key factors:

- **Latency and Time:** This comes down to immediacy. How frequently do we need to collect data, and, more importantly, transform that data into insights for the business process? Data collection and analysis can vary greatly from months, weeks, days, hours, to seconds (even milliseconds). In some cases, companies may be using sensor data to physically control "things" such as equipment within a power plant (e.g., certain readings may reveal the need to adjust settings on a device). In this case, data needs to be collected, analyzed, and actuated quickly. Such efficient data analysis may require compute power closer to the edge of IT/OT to reduce the inherent latency introduced by sending information back to a data center or cloud computing environment. The use of data aggregation and gateways at the edge is a common approach in IoT environments.
- **Bandwidth:** How much and what type of data is being captured have significant implications on bandwidth. Sending large amounts of data, such as images and/or video, to the cloud or a data center could require connection speeds that are either unavailable in the OT environment or cost-prohibitive. Compute resources closer to the edge can compile and filter out data to alleviate bandwidth challenges.
- **Quality:** Some analyses can be quick and simple, such as a binary response: Is a valve opened or closed? On the other hand, companies may want to analyze a much larger data set: perhaps the long-term operating and financial performance of equipment throughout an industrial environment. Quality is often a tradeoff between depth of analysis/insights versus the immediacy of those insights. If depth of analytics is required, but immediacy of the insights isn't critical, then computing resources in the data center or cloud might be more appropriate. Over time, as IoT initiatives become more sophisticated, companies may find increasing need for both depth of quality and immediacy, leading toward more powerful compute resources moving to the edge.
- **Security and Compliance:** By its very nature, security in an IoT system is distributed. Security therefore requires an end-to-end approach that incorporates devices, information exchange, management and control, data distribution, storage, and networks. Connecting a multitude of new endpoints to a variety of different networks yields an expanded attack surface area, and therefore requires some security resources at the edge of the network. Additionally, the connection of mission-critical OT devices to the network raises new cybersecurity threats.

Compliance plays a part here too: Industry regulations and/or local laws may require data to reside in a specific geographical area and prohibit data transfer in some cases.

Depending upon the specific use case, where compute happens will vary, and will often require a distributed architecture. Compute can be encapsulated in devices at the edge, in gateways, and in data centers (see Figure 2). Some demands will be persistent while some will be temporary, and resource allocation needs to be optimized for efficiency while still being able to accommodate peak loads. In big data clusters, processing is often bundled tightly with memory and storage, and an appropriate balance needs to be found to meet the demands of the system. This balance needs to be struck both in the data center and the cloud, but it is also increasingly needed at the edge—or where IT meets OT.

FIGURE 2. Internet of Things Network Data Flow



Source: Enterprise Strategy Group, 2016

Making the Right Decisions for IoT Compute Resources

In an IoT environment, computation needs to be performed as quickly as the associated system of things requires. Some processes, like balancing the energy production loads and consumption on a power grid, will need instantaneous compute, at risk of blackouts and equipment damage. Other uses, like reporting the daily or monthly production totals of a particular assembly line, aren't as immediately time-sensitive. For data analytics, data movement, and data management, all IoT processing needs to be performed at the optimal point across all sites. By optimizing where compute resources exist, organizations can conduct data aggregation, processing, and actualization where—and when—they are needed.

Overall computation power and performance clearly matter, but so do other factors. From a data standpoint, these other factors include data ingestion capabilities and data protection. For servers being deployed at the edge, form factor increasingly comes into play—in industrial environments (e.g., manufacturing, oil and gas, etc.), compute resources must be ruggedized and hardened to meet the requirements of their OT environment. Energy consumption may be a factor not only from a cost perspective, but also from the standpoint of the power resources available. Finally, manageability and lifecycle are increasingly critical in an IoT environment. The ability to manage compute resources that might be deployed at the edge, far from a central data center, in remote locations will be a necessity in most IoT deployments. And the lifecycle of compute resources will also be expected to closely match that of the industrial equipment in the OT environment.

The Bigger Truth

IoT represents the intersection of IT and OT—or where the digital world meets the physical world. This holds enormous transformational potential to unlock business value in the form of operational efficiencies, better and differentiated customer service, creation of new products and services, and development of new business models. However, IoT initiatives by their very nature create new challenges for businesses because of two distinct factors. First, they are inherently large and distributed, with a vast array of sensors and devices spread out across a physical environment. And secondly, they necessitate the integration of information technology (IT) and operational technology (OT).

Given these truths, enterprises must design IoT initiatives with end-to-end strategies—meaning from the data center through the edge to endpoints—that encompass compute, connectivity, security, analytics, applications, and services. To accomplish this, organizations must acknowledge that concepts such as security, compute, and analytics cannot just be executed centrally, but rather require a more sophisticated approach that places all of these capabilities across the IoT environment, including within corporate data centers, in the cloud, and at the edge.

To effectively harness the power of IoT, organizations must answer the What, Why, How, When, and Where of data requirements, which will also instruct computing resources and architectures. To bridge the gap between IT and OT, technology leaders must work together to understand business requirements up front as they develop their compute strategy. Key considerations for IoT compute will include computational performance and reliability, data analytics and protection, form factor and ruggedization, and lifecycle and manageability. Finally, organizations need to recognize that IoT initiatives will evolve over time and planning for the future is critical. Creating a roadmap that begins with initial, small IoT implementations but also validates the growth needs that might arise will allow companies to architect to future requirements.

IoT initiatives that transform that data into business value require a new way of thinking about IT and OT together, particularly at the edge of the network. The edge, of course, is at the very intersection of IT and OT, and the critical juncture where information from the physical world is transferred to data. The edge is no longer just something to be managed, but instead represents a critical place where data processing near the source can maximize the speed of insights.