

THE COMING DATA AVALANCHE – AND HOW WE’LL HANDLE IT

HOW INCREASES IN DATA AND INTEL’S ACQUISITION OF ALTERA WILL IMPACT THE SEMICONDUCTOR INDUSTRY



IN ASSOCIATION WITH: 

TABLE OF CONTENTS

FOREWORD.....	3
THE COMING DATA AVALANCHE.....	3
THE SHAPE OF THE FUTURE.....	4
FROM DATA TO INTELLIGENCE.....	6
HOW TO HANDLE THE FLOW OF DATA.....	7
THE METAL MEETS THE ROAD.....	11
RISE OF FPGAS LEADS TO NEW PARTNERSHIP.....	15
LOOKING AHEAD.....	16
METHODOLOGY.....	17
ACKNOWLEDGEMENTS.....	17

FOREWORD

The Internet of Things (IoT), the installation of sensors almost everywhere, and increases in Internet connectivity are generating an exponential amount of data. This data requires extensive processing – from the point of origin, through the cloud, and in many cases, into systems that perform analysis – leading to the need for a new approach to hardware.

In 2015, Intel acquired Altera with the goal of improving field-programmable gate array (FPGA) technology that is now an invaluable part of meeting future processing needs. FPGAs are becoming essential in a wide range of components, mainly because they can be configured – and reconfigured – by developers, which gives manufacturers the flexibility to continually update and improve these devices. We'll explore this benefit, as well as others, in this report, which is based on research and interviews with today's leading industry executives.



Bruce Rogers
Chief Insights Officer
Forbes Media LLC

THE COMING DATA AVALANCHE

The signs are clear. The amount of data created today is expanding at an unprecedented rate. By 2020, it's predicted that there will be as many as 50 billion devices connected to the Internet¹. With sensors being embedded in practically everything from clothes, to cars, to buildings, and even to cement, we now have access to an enormous amount of information at our fingertips.

Yet left unchecked, this massive amount of data will become uncontrollable and be rendered useless to enterprises worldwide. So how can organizations navigate through this new world? Forbes Insights recently engaged with Intel to find out. Forbes surveyed more than 300 CTOs and system architects to better understand their current computing environments, their future computing plans, the hardware and software they're currently using, and how they expect to accommodate the future, higher performance requirements that ubiquitous data will bring about. We also interviewed industry leaders positioned at the forefront of this new revolution to get their perspectives on the world today, and the world to come as data rates move from a trickle from a garden hose to a blast from a fire hose.

Today's technologies may appear to be managing the current flow of data, but new structures, new models, and new approaches are emerging. We'll look at some of these approaches to managing data as well as how Intel's previously mentioned acquisition of Altera will impact the future of data management, the semiconductor industry, and system development.

“Data is going to change. The bandwidth of cable will change. The amount of data will change.”

--Richard Mark Soley Ph.D.,
CEO, Industrial Internet
Consortium, Object
Management Group

THE SHAPE OF THE FUTURE

The term ‘Internet of Things’ (IoT) has become something of a buzzword recently. It relates to the use of devices that generate and store data, then share it over the Internet or with other devices. In office environments, for example, IoT is used to determine environmental conditions – heating, cooling or lighting requirements – based on information provided by sensors. The sensors not only detect the current environmental conditions; they also detect whether the area is being occupied. Unlike today’s environmental management systems, future systems can be designed for specific zones – with resources applied only to areas requiring them. Such intelligent systems will act on data generated by the sensors, in effect opening the doors to energy efficiency.

Early adopters of new technologies have installed multiple IoT devices in their homes. The Ring doorbell, for example, is a video monitor that incorporates a doorbell, camera, microphone, speaker, and a WiFi device. Ring can detect when a person approaches the doorbell and generate a signal that is transmitted to the owner’s cell phone. The owner can then ‘answer’ the ring, and talk to the person at the door, from wherever he or she is located.

IoT is already impacting all industries from automotive, to wireless communications, to manufacturing and distribution. Local Motors, a multinational technology company that manufactures motor vehicles, for instance, works to improve transportation through 3D printing and robotics technologies. In the past decade, the company has created and launched a number of products, including the Rally Fighter, the world’s first co-created vehicle, which was used to demonstrate IoT capability; the Strati, the first 3D-printed car; and, most recently Olli, the first cognitive, autonomous vehicle to incorporate IBM Watson technology. Olli is the first vehicle to use the cloud-based cognitive computing capability of IBM’s Watson to analyze and learn from high volumes of transportation data, produced by more than 30 sensors embedded throughout the vehicle.

While Olli is the first of its kind, 69% of those surveyed by Forbes insights agree that safe, autonomous, self-driving cars will become a reality in the near future. Creating truly autonomous cars, however, will require the rethinking of hardware. “With autonomous systems, you have to replicate the human intelligence to drive the car,” says Jay Rogers, CEO of Local Motors. “But in order to do that, you have to make higher order computational systems and safety systems,” he says.

Determining how to handle unexpected situations, like how to navigate on a street that has no lane markers, will require a great deal of processing. For a car traveling at freeway speeds, most of the processing will be done on board, and quickly, since there may not be time for the car to generate a message about an unexpected situation and receive instructions. In some cases, it’s conceivable that emergency processes may not be used for basic operation of the car. In an emergency, special instructions would be sent to an FPGA, configuring it to take over the emergency response actions that are required.

Security is another essential element for autonomous vehicles and, in many cases, for much of IoT. In the near future, autonomous cars may be connected to the Internet, sharing relevant data over the cloud, and, possibly, from car to car. The vehicle must be able to apply security algorithms to ensure that the data received is coming from a trusted source. The digital algorithms used for sending data, and for validating and decrypting incoming data, will be complex, and constantly changing.

Like dealing with unexpected situations, handling the security screening will require substantial computing capabilities and FPGAs are ideally suited to performing these tasks. They operate quickly and can be reconfigured when necessary to upgrade security settings.



FROM DATA TO INTELLIGENCE

With so many devices now connected to the Internet, terabytes of data are being generated. Take cellular networks, for example. “We’re seeing an exponential growth in traffic,” says Ron Marquardt, VP of Technology at Sprint. The traffic is mostly consumer traffic, which is less dense than IoT traffic, but the amount of consumer data is steadily increasing. “This means we have to continually evaluate our infrastructure in order to ensure it can sustain an ongoing, exponential increase in the amount of data over the network.”

“There is no point generating an enormous amount of data if we don’t know what to do with it.

--Stephen Mellor, CTO, Industrial Internet Consortium

Faster switching and routing can help manage increasing traffic flows, as can FPGAs. These processors, once implemented in network equipment, can provide higher performance than other silicon devices and also consume less power. This is important for business leaders to consider since, as Forbes Insights discussed in a recent report, *The Mobile Industrial Revolution: Anticipating the Impact and Opportunities of 5G Networks on Business*², more than one-third of executives believe their current systems can’t support the evolving needs of their business. Their companies will require revolutionary increases in the speed, capacity, and connectivity of mobile devices—and they’re looking to new systems, particularly to 5G networks, to provide it.

5G is a communications standard that will open the Internet to ubiquitous computing. “Where 5G becomes important is overall scaling,” Sprint’s Marquardt says. “This will go from hundreds of thousands of devices per square kilometer, from smartphones, to something more ubiquitous. The number of devices may eventually grow to 1 trillion globally, and 5G will factor in as the technology enabling this data growth in the 2020 timeframe.”

This may then lead to the Internet of *Everything*, which will be borne out of several factors. Given that sensors can be developed so cheaply and last for many years, it’s only logical that they’ll increasingly be attached to almost everything in the future.

Pair this with the rise of 5G, which provides data connectivity that can be defined as creating ‘ubiquitous’ connectivity, and almost everything can be connected to one or more sensors and processed by the communications infrastructure.

To take advantage of the insights data can deliver though, companies need to have the right processes and technologies in place.

²<http://www.forbes.com/forbesinsights/huawei/index.html>

HOW TO HANDLE THE FLOW OF DATA

How can companies intelligently handle the flow of data? The assumption is it will be sent to the cloud. “The only place that has enough storage for all this data is the cloud,” says Dean Weber, CTO of storage software developer, Mocana. “Local devices can be quickly overloaded with information.”

Yet this isn’t an opinion endorsed by all. “I’m not a fan of the idea that we’ll send every bit of data to the cloud,” says Dr. Stan Schneider, CEO of Real-Time Innovations (RTI). “Autonomous cars can generate a terabit of data per hour, most of which is feedback control. There are a lot of sensors – video, radar, lidar – and the car needs to process the data right then and there. Although there is certainly value in getting the right data to the analytics, the idea of sending the data to the cloud to control the vehicle doesn’t make any sense. The car has to work even if it’s not connected to the cloud.”

For virtual reality or augmented reality devices, the flow of data from the cloud may have to be very fast so that the visual changes that occur when the user of the device moves his head or body will result in a fast on-device response. This will require high data transfer rates between device and cloud, and rapid processing inside the device or in the processing centers that are responding to movements of the virtual or augmented reality user.



Before data is sent to the cloud though, intelligent processes can determine how the data is managed. These processes include:

Edge Computing

Edge computing puts decision making and intelligence into the devices generating the data. A basic fitness tracker is a good example of an edge device. It stores data, and provides tracking information – either on the device’s screen or through the user’s phone or computer. Edge devices may generate a great deal of data, but they have the intelligence to know when data should be sent to the cloud, when it should be stored, and when other actions or human intervention is required. For example, a 10GB switch in an organization’s computing system will not only perform switching – it will monitor performance, user activity, and other factors. If the switch detects a sudden spike in traffic, it may determine the source of the traffic, apply rules for determining whether this is a denial of service (DOS) attack or transient spike, and may respond to the attack by notifying the IT department and temporarily shutting off the source of the spike. Further, the device will be able to analyze the data using internal security software. All this processing can be done on the edge, and doesn’t require that the data be sent to the cloud.

Fog Computing

Fog computing performs tasks that are similar to those done by edge devices, but operates somewhere between the edge and the cloud. For example, autonomous vehicles handle much of their computing on board, but may also regularly be sharing data, such as road conditions, with other vehicles. The shared information can be used by other vehicles to determine best routes, adjust speeds, and perform other functions, without necessarily involving the cloud. Fog computing can be sensitive to location and can aggregate GPS data, which makes it useful for traffic control and other tasks. Fog computing is also designed to implement security protocols, effectively validating data before it is sent to the cloud. Processing in the fog can be more extensive than that for which edge devices are designed; however, this acts as an effective buffer between the edge and the cloud. Because of the significant amount of processing needed to adequately perform fog services, high speed computing capabilities are required.

Edge and fog computing can involve a number of processes, including:

- **Filtering** – Filtering is knowing what is being processed and should be stored, and discarding irrelevant information. For example, an oil refinery’s tankage system may have hundreds of sensors dispersed over several miles that produce status messages. In the past, a human had to review this data every few hours, but if an emergency arose, problems typically weren’t detected until after they occurred. By connecting the sensors to the fog, the information is sent in cohesive data streams that can help the refinery manage its flow and production. This data can then be used in a data center for refinery control and AI can be applied to enable optimum operational capability at the refinery. Effective filtering, in sum, can restrict the amount of data that should be sent to the cloud from the rest of the data being generated.
- **Aggregation** – Aggregation is the method of aggregating data produced by a device or devices. Light bulbs, for example, produce a lot of data; data that cannot only be filtered, but also aggregated. At preset intervals, aggregate data (1000 light bulbs have burned at full power for the last eight hours) can be transmitted to the cloud or another target. The simple ‘all’s well’ message, summarizing the status of the aggregated devices, spares systems from receiving real-time messages from all the light bulbs.
- **Discarding** – It’s become common belief that anything that goes onto the Internet stays on the Internet. For IoT, however, this belief isn’t true. Much of the data produced by autonomous cars, for instance, is of no value once the vehicle has arrived safely at its destination. When cars operate, they generate, process and filter data, but also store information. Data can be retained for a pre-set amount of time, but it will ultimately not be saved. And it shouldn’t be since, once a car arrives safely, that data no longer has value. If ALL the data that was generated was stored, storage systems would be continually overwhelmed. The intelligent management of data storage – knowing what to keep and what to discard, and perhaps how long and what types of data should be stored – are an essential part of handling data. When intelligence is applied to the types of data being produced, how it’s handled, and where the processing occurs, the amount of data that is actually sent to the cloud can be considerably reduced.

It goes without saying that the processing of data by edge and fog devices requires significant processing capabilities, which makes the use of high-performance, low-latency processing devices essential.

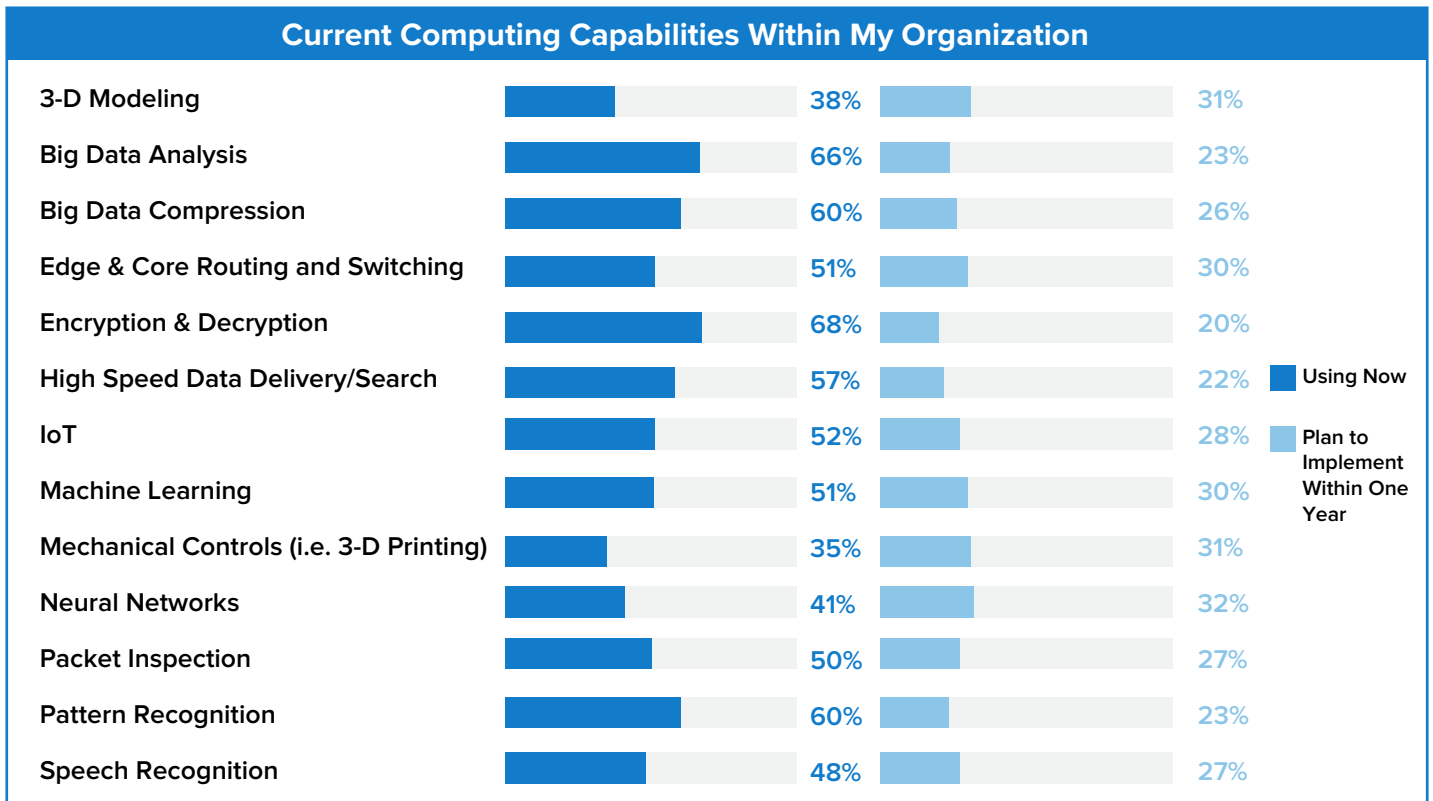
Data may flow in both directions – out of the edge and fog devices, into computing systems where the data is analyzed and decisions made, and back into these devices. Ultimately, instructions may be sent through the cloud, back to the edge devices. Security is an essential factor that must be considered when handling data. This often involves extensive processing to ensure that the user’s security and confidentiality are always preserved as best as possible.

The value of IoT will only be realized when one or more of the following conditions are met:

- Data is unlocked and extracted from sensors and devices
- Data is processed locally on the device. It may also be aggregated for transmission to the Fog or cloud
- Data is shared with business logic IT systems
- Data is analyzed and decisions are made, based on the analysis
- In some cases, responses and instructions are pushed back to the Fog, Edge, or device that generated the data
- In some cases, the instructions may result in the device making changes or sending instructions to other devices. For example, an environmental control system may detect many people arriving at a building after normal hours. After processing this data, instructions may be sent to the lighting and HVAC systems instructing them to turn on.

Processes like the ones listed above (receiving notification of a traffic jam and deciding to send rerouting information to other vehicles) can be applied to many other types of issues. But, at the core, it's a matter of a) learning from the data and b) responding in an appropriate way.

The types of high-level processing described above is already being applied by businesses. For example, out of the organizations that we surveyed, at least 60% are currently using pattern recognition, encryption and decryption, big data compression, and big data analysis. Thirty percent or more of respondents also plan to implement the following within one year: machine learning (artificial intelligence); edge and core routing and switching; neural networks; 3-D modeling; or, mechanical controls (i.e. 3-D printing, robotic surgery). (See Table)

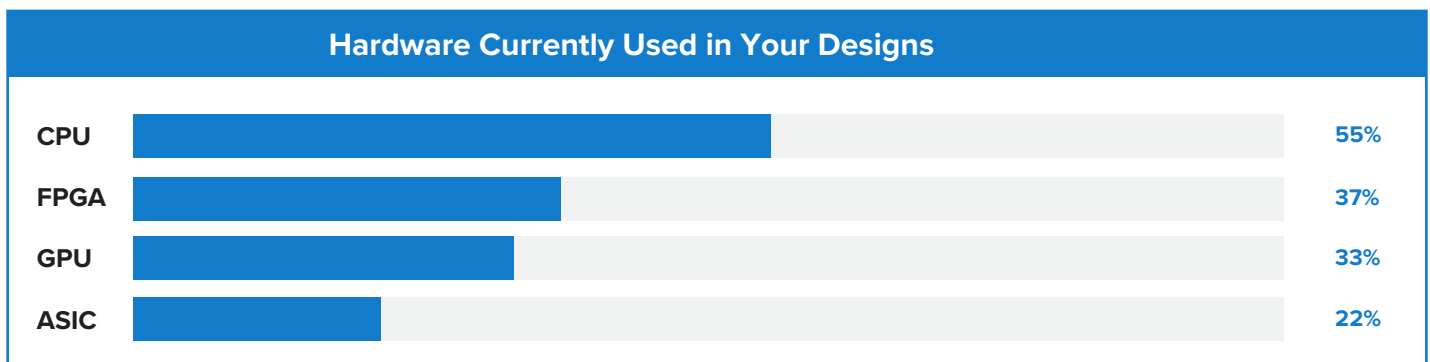


THE METAL MEETS THE ROAD

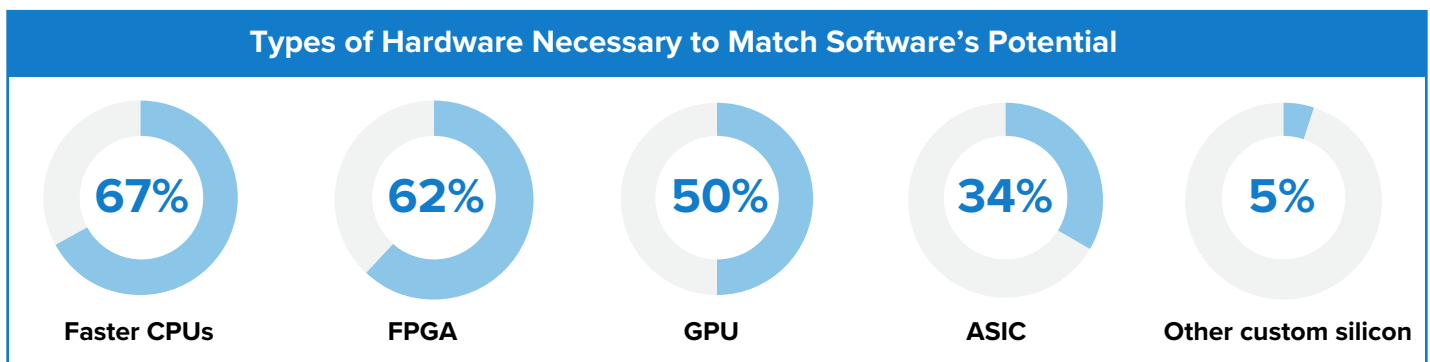
The question now becomes: how will businesses handle this inevitable spike in data flow and the analysis, switching and power management it requires? Current hardware infrastructures are already challenged to keep up with the impending data explosion. Although infrastructures were designed to accommodate expected growth, advances in processing capabilities no longer maintain the improvements they have in the past. That's why a new approach to hardware is needed.

Organizations are currently working to meet future computing requirements. The Forbes Insights survey found that 63% of organizations develop specifications and handle programming internally, while 29% expect vendors to determine hardware requirements and deliver required performance.

So what types of processors, or combinations of processors, are companies using today? Fifty-five percent of CTOs and system architects note that most of their designs use CPUs, but FPGAs, GPUs, and ASICs also play a role when it comes to hardware. (See Table)



How can hardware catch up, and keep up, with software? Sixty-seven percent believe faster CPUs are needed to match software's capabilities, and 62% cite FPGAs. (See Table) There are many reasons why FPGAs were selected as highly as faster CPUs. FPGAs provide performance improvements over other types of processors, and this performance increase can be achieved while using less energy than other processors. FPGAs are also very agile because they can be used in a variety of embedded applications and data center workloads.



“We rely on FPGAs heavily in our equipment. They are a cost-effective way to get processing done.

—Ron Marquardt, VP of Technology, Sprint

For cloud providers, or other organizations doing their own data processing, routing and switching are a major challenge. Microsoft has been developing a hardware infrastructure for its Azure Cloud computing centers. The technology is also being used to accelerate the speed of Internet searches through its Bing search engine. Microsoft's initiative, known as Project Catapult, makes extensive use of FPGAs. "By exploiting the reconfigurable nature of FPGAs, at the server, the Catapult architecture delivers the efficiency and performance of custom hardware without the cost, complexity and risk of deploying fully customized ASICs into the datacenter," says Microsoft's blog. "In doing so, [Microsoft has] achieved an order of magnitude performance gain relative to CPUs with less than a 30 percent cost increase, and no more than a 10 percent power increase."

Microsoft isn't the only company using FPGAs. Wireless providers are also taking advantage of the technology. "We rely on FPGAs heavily in our equipment. They are a cost-effective way to get processing done," says Sprint's Marquardt.

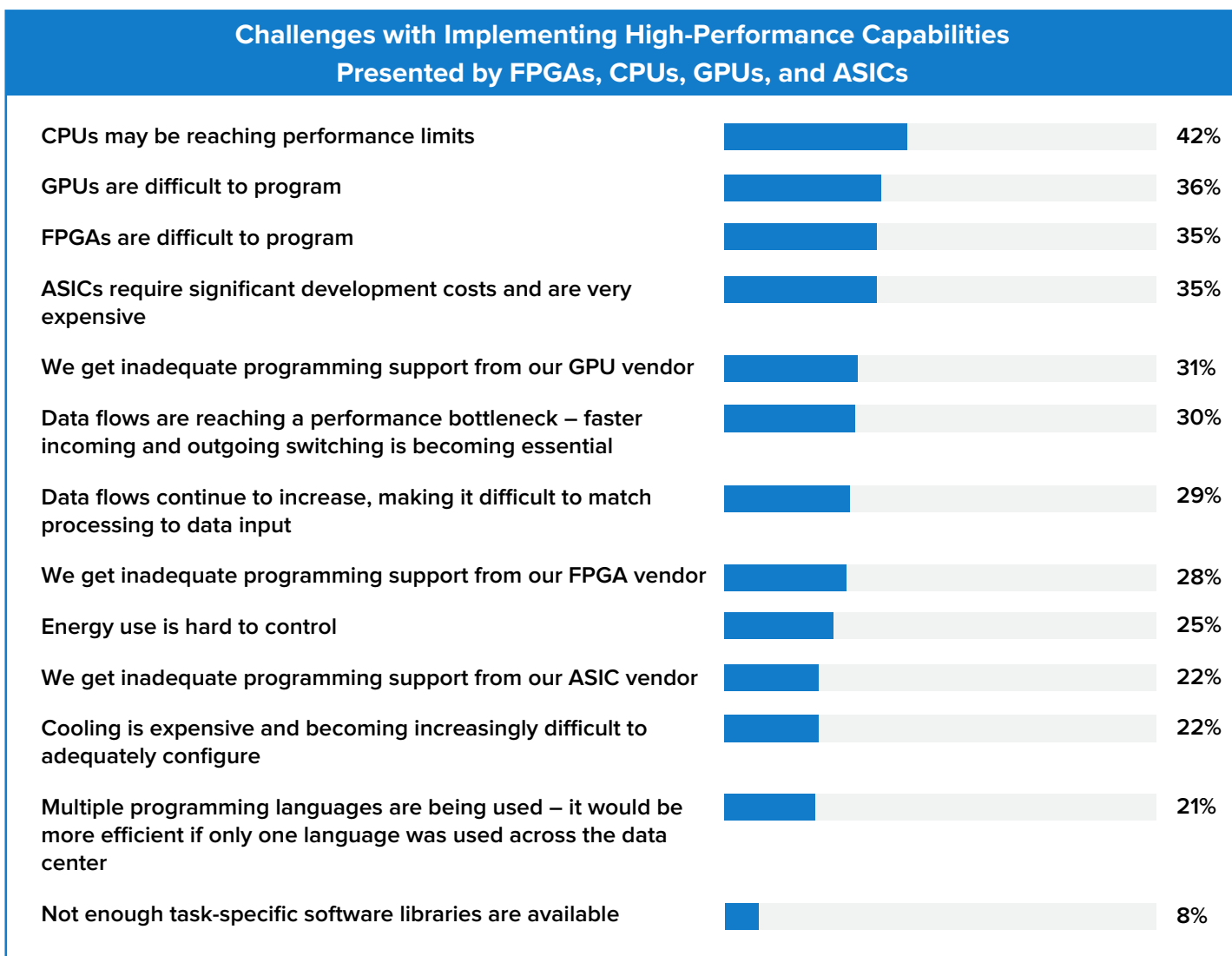
Mocana's Dean Weber notes that some FPGAs will be dedicated to specific tasks, especially in the IoT space. "Security is one of the value propositions that FPGAs bring to the ecosystem," he says. FPGAs can be programmed to process streams of data that use rapidly changing encryption or keys. By rapidly reprogramming FPGAs to process new, dynamically changing security keys, security can become more efficient than it could with silicon, which can't be modified. "One FPGA will be tasked with primary functions, another one with security functions – this approach is still cheaper than CPUs," Weber added. In other words, the use of multiple FPGAs to perform a variety of functions is a better approach than using a CPU dedicated to performing similar functions.

Another area where the easy modification of FPGA operations is coming into play is in robotics. An FPGA can be quickly configured to operate visual sensors, detecting characteristics of the item being inspected, and then reconfigured to control cutting or milling tools. In terms of robotic surgery, an FPGA may be configured to inspect and determine which areas require surgical intervention, and, once this is determined, be reconfigured to control surgical instruments.

In many areas, FPGAs are replacing ASICs because FPGAs allow for hardware and software programmability, significant improvements in time to market, and reduced development costs. “FPGAs were created so that they can have ASIC capabilities, but have some modicum of capabilities for change,” Weber says.

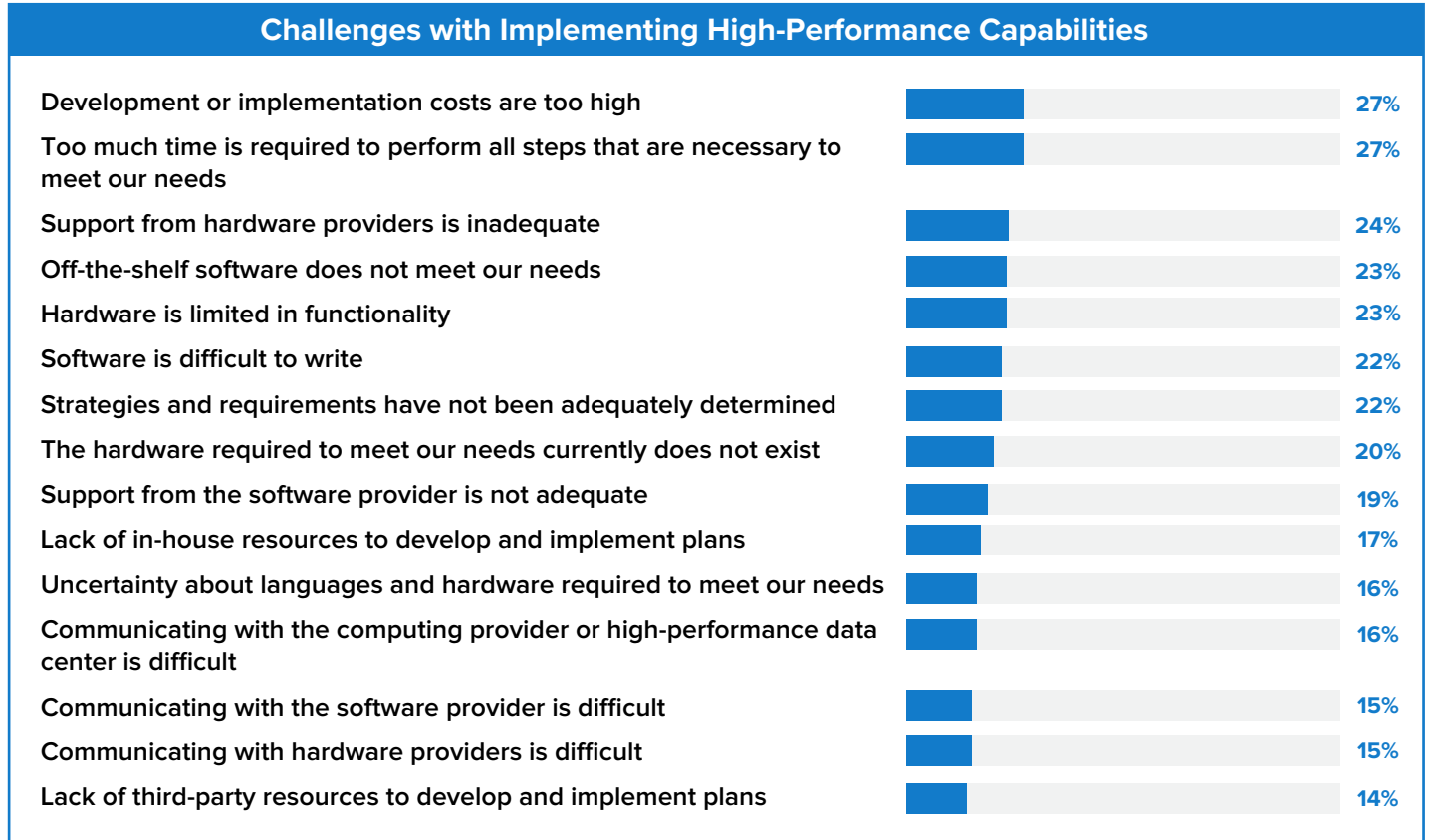
When performance is factored in, FPGAs can actually be an economical option for accelerating many workloads. In cases where a server has to handle a mix of workloads (making ASICs impractical), FPGAs are a very energy-efficient solution, delivering more computations per joule than alternatives.

Yet implementing the hardware backbone for high-performance computing capabilities, while necessary to process and analyze the expected data avalanche, isn’t going to be easy. When working with a variety of processors, CTOs and system architects cite a number of challenges, the top ones being: CPUs reaching performance limits, and GPUs being difficult to program. (See Table)

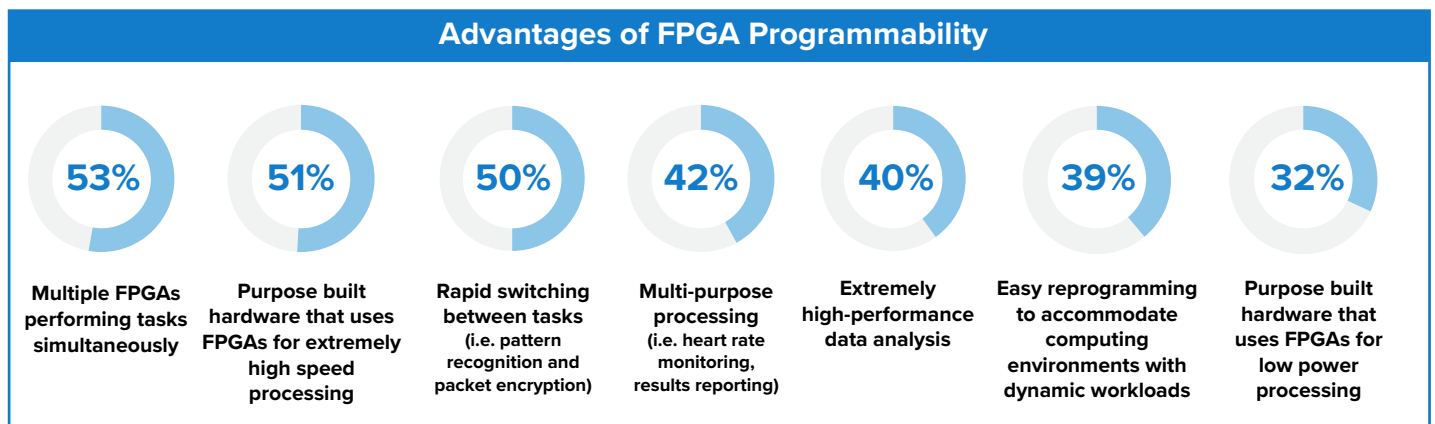


While a third of respondents consider FPGAs difficult to program, this issue is being addressed with the improvement of FPGA programming languages, such as OpenCL, and a broad selection of specially created software libraries. The libraries, developed by Intel and others, simplify the task of programming FPGAs for specific applications.

Even though certain processors present unique challenges to implementing high-performance capabilities, there are obstacles beyond the technology itself, including high development costs, a lack of support from hardware providers, and a likelihood of off-the-shelf hardware not meeting enterprise needs. (See Table)



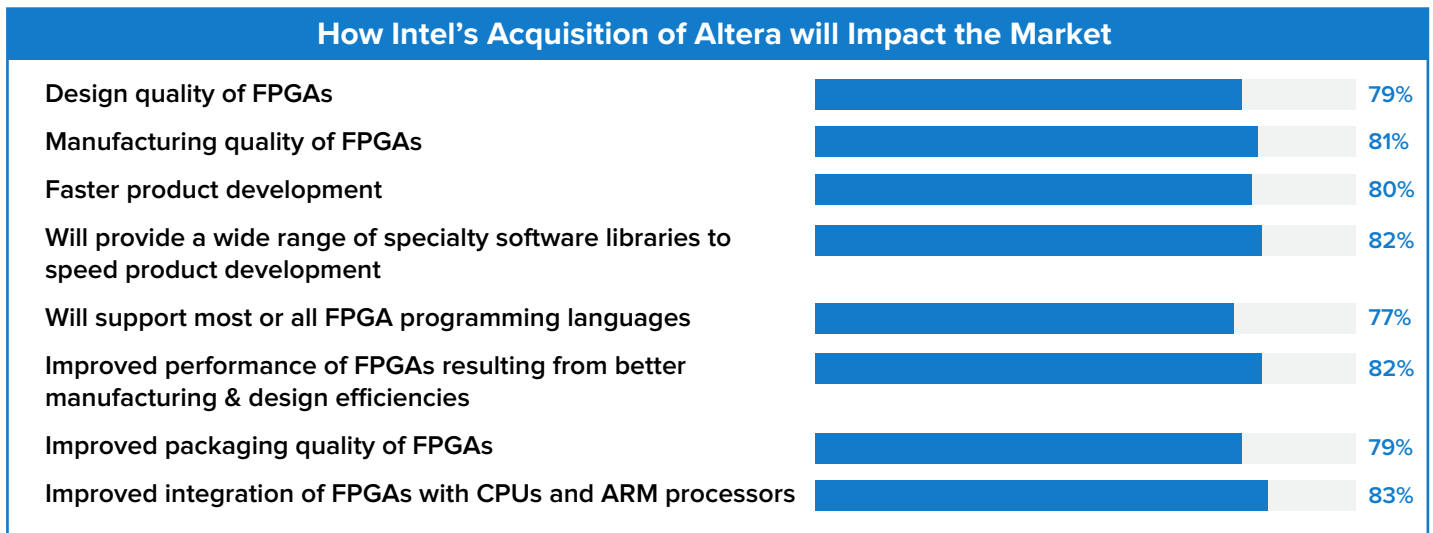
Yet despite all the challenges with implementing high-performance capabilities, in the end, they must be overcome. Organizations that use FPGA programmability in their systems see a range of benefits, including rapid shifting between tasks and multi-purpose processing. (See Table) Microsoft, for example, is taking advantage of an ‘elastic reconfigurable acceleration fabric’ that is achieved by using FPGAs in its servers. “Project Catapult FPGAs talk to each other and in fact, can become pieces of a larger ‘brain’ of an interconnected machine learning network of servers,” according to a Microsoft blog.



RISE OF FPGAS LEADS TO NEW PARTNERSHIP

Given the benefits of FPGAs, in 2015, Intel acquired Altera to encourage new products in the high-growth data center and IoT markets. As stated in an Intel press release, “Combining Altera’s industry-leading FPGA technology and customer support with Intel’s world-class semiconductor manufacturing capabilities will enable customers to create the next generation of electronic systems with unmatched performance and power efficiency.”

How has this acquisition been received by executives around the globe? More than a quarter believe they’ll see faster product development as well as improved design and manufacturing quality of FPGAs. (See Table)



“Intel is thinking of a brand new world,” says Local Motors’ Rogers. “Hardware, especially in cars, is coming along slowly, but FPGAs can be used for low level optimization in a device in order to get the high-speed processing required for autonomous vehicles done. That’s incredible.”

So what does this acquisition mean for your business? With the increased complexity, and amount of data available today, companies need a solution to help them handle this avalanche of data— especially where high performance is required. The products that enable this performance will most likely be FPGAs, or built around them.

In combination with CPUs or other processors, the high-performance and low-energy requirements make FPGAs an attractive choice for integration into products. Additionally, because FPGAs are easily configurable, system designers can change the function of devices into which they’re installed. In effect, FPGAs can become ‘custom silicon’ and can change system behavior dynamically. FPGAs can also be fine-tuned in order to upgrade performance, update security settings, and modify for individual or local needs.

This is something Sprint understands and plans to take advantage of. “The equipment that we’ll purchase in the future will have FPGAs embedded,” Marquardt says. “More and more, we’re relying on (our hardware providers) to give us solutions.”

LOOKING AHEAD

FPGAs, with their high-performance capabilities, low-energy footprint and easy programmability (or re-programmability) will increasingly be an important component in many devices, especially ones around computer security. This is something Intel understands well. As noted above, the company has introduced hybrid devices that combine an Intel® Xeon® processor with an FPGA, and has more hybrid devices scheduled. The idea is that the combination of the multiple processors on the same silicon will enable designers to create products with considerably higher performance than individual devices can achieve. Additionally, Intel is simplifying the interfaces between FPGAs and other system components, and developing programming libraries that will enable developers to more easily program an FPGA, in turn accelerating the development of new devices.

As of today, we've only had a glimpse of this new data-driven world and the computing capabilities it will require. Just about a tenth of executives believe that IoT or machine learning are already here. They do, however, realize that they need to get their organizations ready to handle the enormous amounts of data that will be generated by ubiquitous devices and IoT. Already 65% of IT executives spend at least a quarter of their IT budgets on high-performance computing capabilities. When making your investment decisions, consider:

- Technologies will need to translate this enormous supply of data into intelligence. Without the ability to gather and understand data, and use data-driven insights for smart decision making or actions, there is no point in collecting it in the first place. Translating data into intelligence will require various high-level processing capabilities, including pattern recognition, encryption and decryption, big data compression, and big data analysis.
- The tools that will enable high-performance computing are not yet up to speed. Hardware, especially, is an issue. A majority of CTOs and system architects believe that software currently exceeds the abilities of hardware, which can lead to bottlenecks in data flows. Executives would like faster processing, and this includes CPUs and FPGAs. They are also well aware of the challenges with hardware, such as difficulties with programming and high development and implementation costs. To better handle these challenges, many executives would like more support from their vendors.
- Many systems and operating budgets are limited by power. FPGAs and ASICs offer a superior performance per watt footprint, but only the FPGA is dynamically reconfigurable.

METHODOLOGY

This report is based on a survey of 305 executives conducted in October of 2016 by Forbes Insights. Executives surveyed include 203 CTOs, Heads of IT, and Vice Presidents of IT, and 102 Systems Architects. Thirty-six percent of the executives are based in EMEA, 34% are from the Asia Pacific region, and 29% are in the Americas. The organizations represented operate in a wide range of areas including IoT, knowledge systems, mobile infrastructure, artificial intelligence/machine learning, and cloud services. Executives' companies have at least 250 employees, with 55% having more than 5,000 employees in their organization.

ACKNOWLEDGEMENTS

Forbes Insights and Intel would like to thank the following individuals for their time and expertise:

- Lynda Chin, M.D, Associate Vice Chancellor and Chief Innovation Officer for Health Affairs, University of Texas System
- Ron Marquardt, VP of Technology, Sprint
- Stephen Mellor, CTO, Industrial Internet Consortium
- Jay Rogers, CEO, Local Motors
- Dr. Stan Schneider, CEO, Real-Time Innovations (RTI)
- Richard Mark Soley Ph.D., CEO, Industrial Internet Consortium, Object Management Group
- Dean Weber, CTO, Mocana

Forbes

INSIGHTS

ABOUT FORBES INSIGHTS

Forbes Insights is the strategic research and thought leadership practice of Forbes Media, a global media, branding and technology company whose combined platforms reach nearly 75 million business decision makers worldwide on a monthly basis. By leveraging proprietary databases of senior-level executives in the Forbes community, Forbes Insights conducts research on a wide range of topics to position brands as thought leaders and drive stakeholder engagement. Research findings are delivered through a variety of digital, print and live executions, and amplified across Forbes' social and media platforms.

FORBES INSIGHTS

Bruce Rogers
Chief Insights Officer

Erika Maguire
Director of Programs

Andrea Nishi, Project Manager
Sara Chin, Project Manager

EDITORIAL

Kasia Wandycz Moreno, Director
Hugo S. Moreno, Director
Mark Brownstein, Report Author
Kelly Stahley Crean, Designer

RESEARCH

Ross Gagnon, Director
Kimberly Kurata, Research Analyst

SALES

North America

Brian McLeod, Commercial Director
bmcleod@forbes.com
Matthew Muszala, Manager
William Thompson, Manager

EMEA

Tibor Fuchsel, Manager

APAC

Serene Lee, Executive Director

