How To Navigate the Journey to Autonomous Factories, & Supply Chains

Autonomous systems will require safety, scalable compute, flexibility, reliability and robustness

Why you should read this

Autonomous workloads are changing the way we design products such as autonomous and semi-autonomous vehicles. It's an enormous opportunity to scale design with more flexible, cost-effective methods to deliver products that are more reliable, robust and safe. But there are roadblocks along the journey. Here are the drivers, challenges and potential solutions to enable this vision of the future.

Introduction

One of the most profound opportunities to propel the digital world forward lies in deploying autonomous workloads

They promise to deliver scalable devices and systems that can safely decide and act for themselves to enable operational efficiencies working alongside humans or taking humans out of the loop, creating time that can be spent elsewhere.

Autonomous vehicles are the most visible manifestation of autonomous workloads, but the factory lines and supply chains that deliver them stand to gain just as much. In the factory, autonomy will lead to increasingly more efficient and intelligent robotic machines tied into ever more sophisticated building and networking systems that will deliver gains not only in manufacturing but in the larger supply chain and logistics operations. And autonomy provides the intelligence to customize workflows, to optimize production line logistics and output.

This will not be an overnight transformation; it will be a journey. Autonomous workloads will need fundamental building blocks including heterogeneous compute, safety, security and machine learning, that are common across autonomous-based use cases.

While the journey will transform how we design and deliver powerful new systems and applications, it won't be easy. The sheer complexity of a fully autonomous system is mind-boggling. Despite this, we're already seeing disruptive companies take advantage of the technology in the industrial sector, including new entrants from non-industrial segments, such as Amazon and Google.

In short, realizing a successful journey to autonomous requires a fundamental reconsideration of design, embracing emerging technologies and methodologies.

Here, we outline some challenges and opportunities in the autonomous journey in industrial, based on a number of interviews with domain experts we conducted to get their observations. The following is based on those interviews.

Section 1
What's driving the need for autonomous workloads

For the industrial sector, autonomous workloads are yet another milestone on a long road of technological innovation dating back 200 years to a period when innovators harnessed rivers to power mills.

Today, the need for autonomous decision-making workloads that are not dependent upon human interaction is clear for a number of reasons.

Managers in the industrial sector want to deploy autonomous workloads to achieve higher productivity and improved KPIs and better operational safety. They want to make their factories more flexible (and therefore more profitable) in building customized products more efficiently, adapting their product mix on the fly.

To transform their business, managers want to integrate their factories more tightly with the wider supply chain and logistics providers. Here, the goal is one autonomous chain from ordering, to raw material requisitions, to last-mile delivery, taking humans out of the loop.

In the industrial world, automatic guided vehicles (AGVS) will benefit from becoming fully autonomous, albeit within a bounded environment. This area will benefit from the trickle down of fully autonomous vehicle research and will result in improved efficiency and customization for factories. It will improve manufacturing by delivering fast-moving consumer packaged goods (FMCGs), pharmaceuticals, food and assembly automation from high-speed production lines that will experience minimal or no downtime or reprogramming.

But at present, factories can be outfitted with disparate systems and a dizzying range of software and hardware choices, which can make charting a path to autonomy challenging. Most factories have some degree of automation, but retooling for new products or expanded customer requirements can take operations offline for days or weeks at a time.

Section 2
Rethinking the compute and data challenge

To realize the vision, we need to understand how fundamentally different autonomous workloads are from traditional workloads

Autonomous workloads are defined by receiving input and generating an output in their simplest terms – the ability to sense and take intelligent action based on the captured data. The magnitude of workload is far greater thanks to the incorporation of ML or AI.

Autonomous is more complex than traditional workloads because of the overlapping and interconnected nature of systems in industrial applications. There may be a machine vision system with its own set of data and algorithms that needs to integrate with autonomous mobile robots (AMR) that are alerted by the vision system that a task has been completed and pickup is required.

Integration may also be required for other points along the supply chain. For example, custom trim requests on a vehicle can be relayed from the buyer or dealer up the supply chain to order and ship what's required.

In industrial machines, there are the enormous computational requirements around capturing, fusing, analyzing and acting on vast (and different – cameras and computer vision to GPS inputs) amounts of data. Layers of sensor inputs must be processed in real time in order to predict situations that machines can anticipate and react to. The technology is there, but the computational power to run the algorithms that process the data is what makes the difference. That difference is defined by the SoC responsible for processing, planning, and control workloads. This all requires a level of redundancy that is not present in other, more traditional workloads. It can be managed efficiently by a single, scalable architecture, however.

Configuring these systems is costly. Compute alone can cost thousands of dollars, while the sensors themselves sometimes cost thousands of dollars. Factory safety demands redundancy across the range of cameras and sensors, and that raises costs. And in general, the more autonomy, the more sophisticated the system, the more expensive it is.

Section 3
How to achieve scale

The real value lies not just in delivering autonomy to devices but by scaling that to a larger context, be it a factory, a supply chain or, in the case of autonomous vehicles, an urban environment.

In the factory, for example, it's not just the autonomy of one robot, but the collection of smart devices within the factory that's required to autonomously decide what to build next and how to do it. This requires orchestration of a collection of systems in tiers of hierarchical autonomy all working together. For example, autonomous robots deliver raw materials or sub-assemblies to an autonomous line. This all requires the right level of compute within the device and system to ensure efficiency, accuracy, reliability and safety.

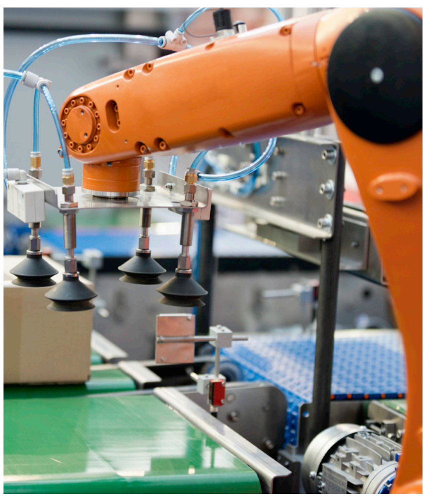
In trying to deliver on this vision of scalable autonomy, many adopters struggle with a high degree of manual work that's required up front to deliver systems. This includes the cost and complexity of silicon and algorithms. Then there's time-to-yield if: In a factory, if the product being produced will command high margins, then the line can afford to be down for six months while it's being made autonomous. But that's not possible for low-margin or low-revenue product lines; manual workflows will still be required.

In addition, engineers worry about the current state of AI capabilities – deep learning and affordable vision and sensing technologies – are not yet where they need to be to deliver the results they want.

Today's approaches and solutions often require developers to bundle off-the-shelf hardware just to get the horsepower required for prototyping. Going forward, these new systems must leverage heterogeneous compute not only to deliver flexibility, but to help reduce the power, size, cost and thermal properties while retaining the performance needed to run autonomous workloads. Heterogeneous compute will be delivered for different stages in the workload and include security enclaves and safety islands to deliver trust.

Section 4
Standard software that scales

For autonomous systems to succeed, there's an entire software stack and software ecosystem that needs to be developed that abstracts many of the details about how the underlying systems work from the people who are thinking about the application level.



For example, it's not going to make sense for every single program that runs on a given piece of hardware to understand exactly what that the interface is to its network accelerator and exactly how to communicate with it. "Hiding" that beneath a standard, and leveraging a robust, open ecosystem with vast number of contributors will help scale autonomous systems.

Traditional companies offer "sticky" proprietary software with no room for the developer to put his or her own programs on. Newer entrants have given these stalwarts a wakeup call with open systems, enabling the developer to have access to it or at least some portion of the control system to put their own control. Autonomy needs to balance openness without losing all "stickiness" in order to head off the disruptors who propose totally open systems. A robust ecosystem can ensure this as well as develop a frame to ease technology selection that helps match an algorithm to a range of hardware with ease.

Customers that have much more demanding safety requirements will require safety-certified RTOS and middleware solutions. All of these elements will have some very demanding requirements as far as being able to achieve safety certifications such as IEC61508.

The software portion of the design also depends on where you are in your development. For example, as autonomous vehicles move from prototype to production or development to deployment, there will be the need for commercial software stacks offering capabilities like real-time, fault tolerant, safety and security, and hence the commercial software ecosystem will be needed to replace of augment the open source and proprietary software that is currently in development.

It's critical to consider some common base architecture that allows your design team to pivot from open source to commercial SW as they reach deployment. This reduces complexity and improves time to market through quick migration of workloads. Software also needs to scale across verticals, so many companies can achieve efficiencies as they design for different applications.

Section 5
The 360-degree approach

To deliver on this vision and overcome the complexity issues, we have to not only bring down the cost of compute but also deliver choice and flexibility to developers ideally with a support ecosystem that can deliver everything from IP to algorithms to test and validation programs and protocols.

New technologies in the design of CPUs, GPUs and ISPs will offer the crucial steppingstones to enhance capabilities in factories and the supply chain to enable more intelligent, configurable production lines in factories as we move into this new era of autonomous decision making.

CPUs need to offer the ability to run different workloads/application footprints, allowing the deployment of the same SoC compute architecture into different domain controllers. They need to support features to help to achieve both SIL2 and SIL3 safety requirements, allowing flexibility through its partitioning features that enable the highest safety levels when they're needed, or highest performance when they are not.

GPU technology needs to deliver "flexible partitioning" to allow the allocation of hardware resources at runtime – whether it's between safe and non-safe workloads or different types of workload. Such technology must be designed to IEC61508 and be SIL2 safety capable, allowing the use of GPU resources for safe heterogenous compute needed in autonomous systems.

Keeping with heterogeneity, developers need an image signal processor (ISP) that can scale to support virtual cameras of varying types and offers a wide range of data output formats providing the flexibility to support both human and machine vision applications such as production line monitoring and quality control.

Conclusion

The autonomous development occurring today and in the future will transform the world as we know it.

All the hardware required for autonomous systems needs to be considered in a comprehensive way to run the complex workloads that are needed. Heterogeneous and scalable compute will address the range of processing capability for differing workloads in a power efficient manner. Enabling this more efficiently and cost-effectively will be standardized chipsets, geared to specific purposes, that are available from various vendors to give designers far greater freedom of choice, flexibility and peace of mind.

For smart manufacturing, autonomy provides the intelligence to customize workflows, whether that's deciding what to manufacture next, or to optimize production line logistics. Factory production lines will become reconfigurable, enabling flexibility to meet the needs of differing products and their specific customisations, and this will add value and save cost. The larger supply chain will become much more efficient and deliver products more quickly thanks to this technological transformation.

Harnessing waterways spawned the industrial revolution, which transformed the world. Autonomous workloads, delivered in the manner described here, will deliver no less profound an impact on the industrial sector and civilization