WHITEPAPER

How AI Edge Platforms Can Transform Smart Spaces

Published by

MOBILE WORLD LIVE

Executive summary

Artificial intelligence (AI) has the ability to transform the way cities and enterprises operate. From traffic management to street lights to retail stores and manufacturing plants, AI can streamline and optimize the way data is collected and analyzed.

In a smart space environment that can mean any location that is equipped with networked sensors that provide its occupants with detailed information about how the space operates or any type of technology that promotes connectedness and creates a more collaborative, interactive environment.

AI-driven tools are particularly powerful when combined with edge computing in a smart space because they create an intelligent edge that enables data to be collected and processed closer to endpoint devices. This keeps traffic local and therefore reduces latency and improves security. It also moves key applications and servers closer to the people and devices that use them.

The intelligent Edge is expanding rapidly. Consulting firm Deloitte predicts that in 2021, the global market for the intelligent edge will expand to $12 billion and continue to grow at a compounded annual growth rate of 35%. The firm also predicts that by 2023 approximately 70% of enterprises will run some amount of data processing at the edge.

Enterprises view the edge and the Internet of Things (IoT) as very important components of their business. In fact, according to a GSMA Intelligence report, "IoT Revenue: State of the Market 2020," enterprises are deploying IoT for three reasons — to achieve cost savings, to generate new revenue and to comply with regulations. Because of this, many enterprise deployments, at least initially, are focused on applications that offer operational efficiencies such as quality control, and management of supply chains, assets and fleets.

The GSMA also found in a survey that it conducted with enterprise IT departments that many enterprises, about one-third of those surveyed, are already using AI and machine learning to make sense of their data. Interestingly, when asked what they do with the data generated by IoT devices, 52% of those enterprises surveyed said that they are using it for edge computing and 51% are using it for performance visibility. About one-third, or 33%, are using that data to create new revenue.

These insights are important because they show that AI-enabled edge platforms are already starting to be used by enterprises and also provide some visibility into the potential that they have to transform enterprises, particularly in their physical environments, which include smart buildings, connected manufacturing plants, retail stores and hospitals.

In this whitepaper, we will explore how AI-enabled edge computing platforms are reshaping the way cities, enterprises and venues operate. Mobile operators play a key role in this journey as they not only provide the underlying connectivity but also bring together the various stakeholders necessary to make this transformation.

Introduction: Evolution of Smart Spaces

Smart spaces are becoming increasingly prevalent all around us. Loosely defined, a smart space is any location equipped with a network of sensors that can provide occupants or owners with better information on how the space operates. A smart space is a term that is broader than a smart city and is often used to describe stadiums, office buildings, factories, hotels, retail stores and even manufacturing facilities. Smart space IoT solutions include traffic management, waste management and street lights.

Research firm MarketsandMarkets estimates that the smart space market will grow from a $9.4 billion market in 2020 to a $15.3 billion market by 2025 with a compound annual growth rate of 10.2%. The firm cited increasing green building initiatives and environmental concerns with driving the growth in this area.

Smart spaces take advantage of IoT technology by bringing together hundreds or thousands of embedded Internet-enabled sensors. These sensors provide a rich set of data that companies can then analyze and use to track assets, gather information about the safety of their operations, or even to automate processes and reduce their labor costs. IoT data can also be used to find out about people's preferences and behavior.

The basic elements of an IoT system include the devices that are used to gather data. Those devices range from the complex, such as autonomous forklifts that move equipment around factories, to the simple such as sensors that monitor the temperature inside buildings. The data from those devices is then collected, processed and analyzed. It can be processed locally at the network edge or in the cloud or a data center.

Here are a few examples of smart spaces:

- An office where the thermostat automatically detects you are there and adjusts the temperature based upon your preference.
- A retail store that automatically detects when you enter, knows what you put in your shopping cart and automatically charges you.
- A factory where machines send and receive information from other machines without human intervention.

AI-Powered Services and Applications for Smart Spaces

Not only do smart spaces use technology and connectivity to make enterprises operate more efficiently but they also enhance daily life and can help reduce the environmental impact on the earth by conserving natural resources and promoting sustainability.

By 2050 it is estimated that 70% of the world's population will be living in cities. This rapid urbanization will create additional challenges in terms of traffic congestion, energy consumption and the associated health risks from carbon emissions and pollution.

But it is possible to create more sustainable cities if governments, municipalities and businesses use technologies like AI and IoT to improve the quality of life. These technologies can be used for many things, including enhancing energy efficiency and waste management, improving air quality and optimizing traffic flow and safety.

We're already starting to see early implementations of these technologies that are helping to create more sustainable cities and communities. Here are some examples:

- In Sao Paulo, Brazil, Telefonica Brazil has developed a way to predict air quality using AI and analytics. Data is collected from weather, traffic and pollution sensors and then aggregated and analyzed. This helps the city calculate and estimate pollution levels up to 24 to 48 hours in advance so decisions can be made to prevent unhealthy air quality levels.
- In Delhi the government has installed an intelligent traffic control system that includes CCTV cameras, automated traffic lights and sensors. All these devices carry and compile real-time data on traffic conditions. AI is then used to process that data so that local authorities can make decisions on how to balance traffic flow and identify traffic patterns and congestion trends. This helps city managers and traffic managers create long-term plans to alleviate traffic problems.
- In Dallas, AT&T Stadium is outfitted with 5G, giving fans access to 3D video and augmented reality that can be processed at the network edge. But there are many more applications that IoT can enable at sports stadiums. Because stadiums are almost like small cities, IoT and edge computing could be used to collect data on the stadium's occupancy rate to make quick decisions about capacity or monitor lines of fans as they enter the stadium to make sure they are keeping six-feet apart. IoT solutions could also be used to solve traffic and parking problems that may exist outside the stadium.
- IoT also can be a game-changer when it comes to disaster management. Real-time sensor data can be used to monitor water levels in reservoirs or detect wildfires, tornadoes, earthquakes and other natural disasters. Perhaps even more relevant is how IoT could be used in a pandemic like what the world is experiencing right now. AI and edge computing are used in healthcare situations to remotely monitor patients. Sensors could be used to make sure people in indoor settings are keeping the required distance apart from each other, or video could be used to make sure face masks are being utilized and any other public health guidelines are being followed.

Supermicro Outdoor Edge Server

Video's role in smart space applications

Although much of the discussion surrounding IoT has to do with low-tech sensors, video actually plays a pivotal role in IoT because it allows users to take a much more proactive approach to business and operations. However, the amount of data created by video is enormous and is expected to grow exponentially over the next few years.

A July 2020 forecast from IDC estimates that the worldwide smart video market will grow to $44 billion by 2025, up from $23.6 billion in 2019, with a five-year compound annual growth rate of nearly 13%.

What distinguishes a smart camera from video cameras of the past is that they have the capability of analyzing video in real-time and, in some cases, they can even see things that aren't apparent to the human eye. Smart camera content can also be trained to do more effective real-time analysis and identify situations and behaviors with high confidence. In some cases, smart video provides critical assistance to public safety officials in reducing crime and improving operational efficiencies.

Here are a few examples of how video is being used in US smart space applications today:

- Bay Area Rapid Transit (BART) was losing over $25 million per year due to passengers evading paying fares. BART repurposed its security cameras to study consumer behavior and was able to improve access and makes some shifts in enforcement personnel. The result was better control and a better customer experience with non-invasive technology.
- Retailer Timberland used NVIDIA's SkyREC AI monitoring and shopper analytics to gather intelligence about what areas of its stores were most popular with customers, and also learn more about customer demographics. Based upon what they learned they were able to more precisely market to customers, resulting in a 30% increase in sales.
- Rolloos Oil & Gas used a network of cameras and sensors combined with deep learning technology to monitor workers on oil rigs and determine if they are in unsafe locations and alert them before an accident occurs.

Another compelling application is using video for roadway tolling. By using real-time analysis of video, transportation departments can determine what type of vehicle is using the road so they can determine the appropriate fee.

However, this type of real-time analysis requires massive computational power. Historically, the data captured from these systems was extracted and analyzed afterwards in a manual process. But today by using AI with a trained neural network it is possible to provide sufficient computing capacity at the point of data capture, which is often referred to as 'inference at the edge.'

This inference at the edge is beneficial because it makes the data more secure as it is stored locally rather than sent to the cloud. Another advantage of the edge is that it reduces latency and overall network congestion.

Today, with the right type of engineering, it is possible to integrate video into smart space applications and make sense of the subsequent flood of data. Companies should be confident that they can create, deploy and scale AI and IoT applications from the edge to the cloud.

Metropolis Platform for Edge AI

Video is one of the largest generators of data and the only way to process that data is by using intelligent video analytics tools that can use deep learning in cameras, on-premises video recorders and servers to monitor and process the data quickly and accurately.

NVIDIA Metropolis is a video analytics platform that applies deep learning AI to video streams. The platform allows companies to use NVIDIA GPUs to create solutions for video capture, processing, storage, and visualization from the edge to the cloud.

Metropolis incorporates many NVIDIA products into one unified architecture. High-performance deep learning inferencing at the edge is handled with NVIDIA® Jetson™ embedded computing platform and through servers and data centers with NVIDIAs GPU.

Rich data visualization is powered by NVIDIA Quadro professional graphics and the entire edge-to-cloud platform is supported by NVIDIA's software stack including DeepStream SDK, Transfer Learning Toolkit, and TensorRT.

Enabling AI at the Edge

GPUs, once exclusively found in high-performance PCs and workstations, now power deep learning training and inferencing in data centers worldwide. Yet low-latency requirements of edge applications require local, powerful processing. Standard data center equipment is not built for Edge applications where power and space limitations are common and environmental conditions demand more robust devices.

Supermicro 1U 4-GPU Server

Supermicro supplies a growing portfolio of servers that can support a range of edge AI workloads with the right CPUs and GPU accelerators and operate efficiently in virtually any needed location, indoor or outdoor. These can range from industrial gateway devices and IP-65 pole-mounted outdoor servers to rackmount servers that are physically optimized for telco hardware and environmental specifications such as NEBS Level 3 while providing data center-class performance.

NVIDIA-Certified Systems™ and ecosystem

To help make it easier and faster for customers to use NVIDIA's latest GPUs, the company collaborated with top OEMs around the world to drive AI forward across every industry. NVIDIA-Certified Systems™ consist of NVIDIA GPUs and NVIDIA networking installed in an enterprise-class server.

Supermicro's servers are NVIDIA-Certified and shipping today. That means that Supermicro's servers have passed a set of certification tests that validate the best system configurations for performance, manageability, scalability, and security.

The Metropolis platform also comes with a partner certification program that offers system design tools and software that help developers build enterprise-ready intelligent video analytics solutions. The certified software, once deployed in the field, makes it easier for companies to manage and upgrade software and scale their video analytics solutions. The platform also includes high-end smart sensors as well as intelligent appliances, such as network video recorders, for on-premises deployment and servers in the cloud.

Supermicro 4U 10-GPU Server

Conclusions

Over the past few years, it's become increasingly clear that AI has the ability to transform many aspects of life — from how we travel, to how we work, and to how we shop, AI can improve efficiency and safety through the analysis of the data we generate.

AI-driven tools are particularly powerful when combined with edge computing and high-throughput 5G connectivity because that enables data to be collected and analyzed in real-time. It also moves key applications and servers closer to the people and devices that use them.

Mobile operators play a key role in edge computing and are also critical to bringing all the various stakeholders together —whether that's municipalities, tech companies, platform creators or application developments — the underlying 5G connectivity will provide a critical element to making AI-driven smart space initiatives a reality.

As a global leader in high performance, high efficiency server technology and innovation, we develop and provide end-to-end green computing solutions to the data center, cloud computing, enterprise IT, big data, HPC, and embedded markets. Our Building Block Solutions® approach allows us to provide a broad range of SKUs, and enables us to build and deliver application-optimized solutions based upon your requirements.

Learn more at www.supermicro.com/iot-edge

Produced by the mobile industry for the mobile industry, Mobile World Live is the leading multimedia resource that keeps mobile professionals on top of the news and issues shaping the market. It offers daily breaking news from around the globe. Exclusive video interviews with business leaders and event reports provide comprehensive insight into the latest developments and key issues. All enhanced by incisive analysis from our team of expert commentators. Our responsive website design ensures the best reading experience on any device so readers can keep up-to-date wherever they are.

We also publish five regular eNewsletters to keep the mobile industry up-to-speed: The Mobile World Live Daily, plus weekly newsletters on Mobile Apps, Asia, Mobile Devices and Mobile Money.

What's more, Mobile World Live produces webinars, the Show Daily publications for all GSMA events and Mobile World Live TV – the award-winning broadcast service of Mobile World Congress and exclusive home to all GSMA event keynote presentations.

Find out more www.mobileworldlive.com

Disclaimer: The views and opinions expressed in this whitepaper are those of the authors and do not necessarily reflect the official policy or position of the GSMA or its subsidiaries.