Solution Brief



Enabling AI at the Edge with Intel® OpenVINO™

Industrial grade computing hardware featuring Intel 13th Gen Core Processing, paired with Intel Arc GPUs, enables rich and flexible AI solutions at the edge using OpenVINO.

Executive Summary

This brief explores how leveraging the Intel OpenVINO toolkit at the edge empowers organizations to unlock the true potential of AI applications. We'll look at how technology innovators can take full advantage of the data they're already generating by combining OpenVINO, Intel's suite of Edge AI Reference Kits, and reliable edge hardware.



Challenge: Limitations of cloud-based AI

While cloud-based AI has revolutionized various industries, its limitations become apparent when attempting to process the vast amount of data generated at the network's edge. Current approaches can hinder effective edge AI applications due to:

Latency and Bandwidth Constraints: Relying solely on cloud-based AI processing creates latency issues, especially when paired with bandwidth limitations. This delays insights and hinders real-time decision-making at the edge.

Security and Privacy Concerns: Sending sensitive data to the cloud can increase the risk of data vulnerability. In addition, regulatory compliance related to data management can become a prohibitive concern for large scale deployments.

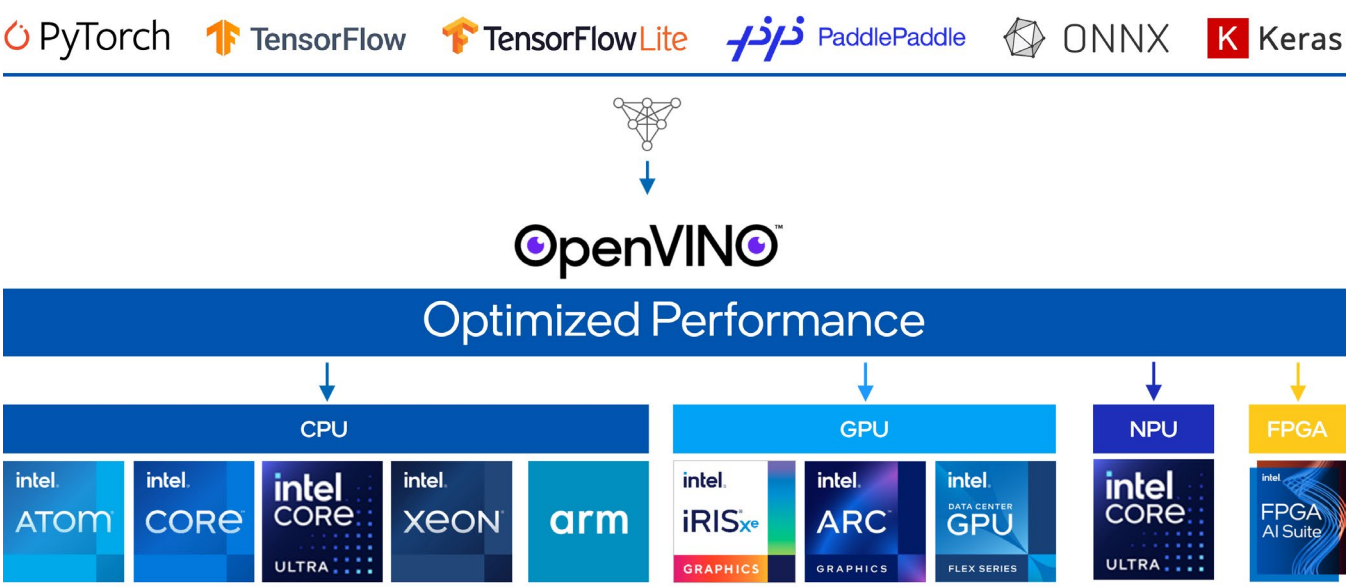
Scalability and Cost: Cloud-based processing incurs ongoing costs as data volume increases. Scaling cloud resources can be complex and expensive.

Complexity: AI deployments are inherently complicated, so the fewer variables in the process, the better. Keeping data in-house by deploying AI at the edge removes a significant part from that complex equation.

Solution: Intel OpenVINO at the Edge

OpenVINO is a powerful toolkit designed to optimize and deploy deep learning inference models on various hardware platforms, including edge devices. The OpenVINO toolkit can help streamline and simplify complex AI models so they can be run locally, rather than requiring vast amounts of cloud computing resources.

Imagine you have a complex recipe (AI workload) that requires specialized equipment to execute (like cloud computing). OpenVINO helps to translate that recipe so it can be prepared using a wider range of tools (like an edge device). This allows you to make the same dish (the AI application), right in your own kitchen (at your network's edge).



The Intel OpenVINO toolkit can be used to optimize and deploy AI models across a mix of Intel hardware and environments.

By deploying AI processing at the edge, you can:

Reduce Latency

Real-time data analysis and decision-making at the source, leading to faster response times and improved performance.

Enhance Security

Sensitive data remains on-premise, minimizing security risks and ensuring regulatory compliance.

Improve Scalability

Leverage existing edge hardware for AI processing, reducing reliance on expensive cloud resources. OpenVINO simplifies model deployment across diverse edge devices.

Lower Costs

Reduced reliance on cloud processing translates into significant cost savings in the long run.

Benefits for Your Business: Speed, Efficiency, and ROI

Deploying Intel OpenVINO at the edge goes beyond just technical advantages. By processing data and making decisions locally, your organization unlocks a range of strategic benefits that will continue to pay dividends over time. AI deployments powered by OpenVINO can help you to achieve:

- **Faster Time to Action:** Real-time insights from edge-based AI enable quicker decision-making and improved operational efficiency, leveraging OpenVINO's "write once, run anywhere" flexibility.
- **Enhanced Operational Efficiency:** Automate tasks, optimize processes, and predict maintenance needs at the edge, leading to significant cost reductions.
- **Improved Customer Experience:** Enable features like real-time product recommendations, personalized experiences, and proactive customer service at the edge.
- **New Revenue Streams:** Develop innovative AI-powered solutions and services that unlock new revenue streams and then feed them back through the process to further optimize.

Next Steps

Every journey toward deploying effective AI solutions is unique, but broadly there are a few key next steps to consider when you're ready to explore how you can best use OpenVINO at the edge, including:

- **Detailed Needs Analysis:** The first step to a successful project is a vision for where you want to go. Be clear in your objectives, what you hope to achieve, what you plan to measure and how you'll know things are working. This step will guide everything from architecture design and ecosystem partner selection, to hardware right-sizing and deployment plans.
- **Proof of Concept (POC):** Partner with experts like Intel and OnLogic to develop a POC that showcases how OpenVINO can be used for your particular AI workload.
- **Scalability Roadmap:** Develop a comprehensive plan for deploying OpenVINO AI workloads across your edge infrastructure, ensuring it can grow with you and drive future success.



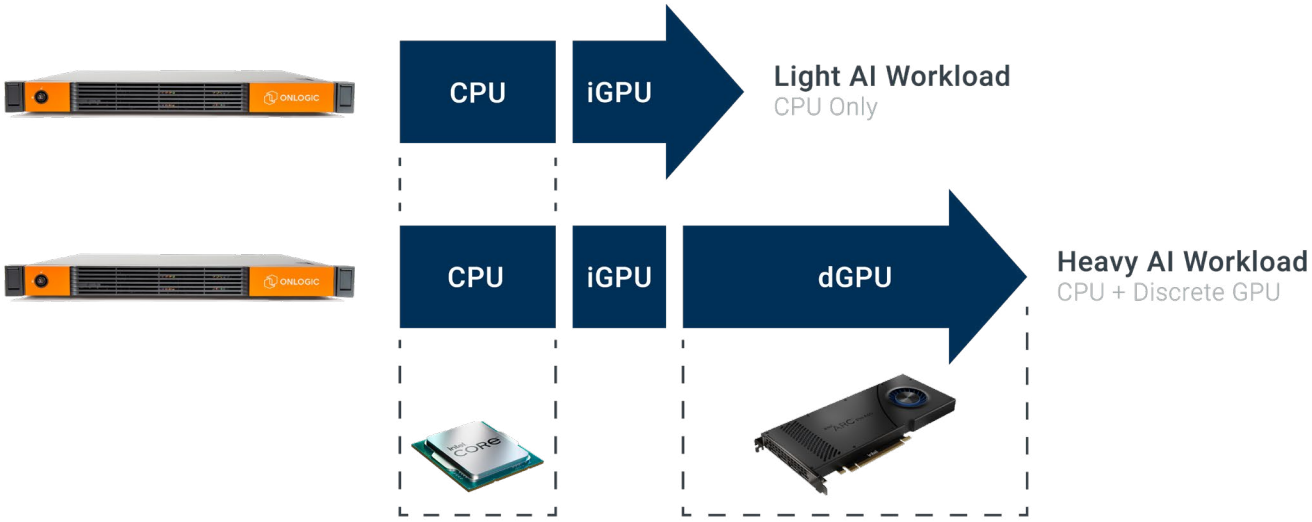
Effective OpenVINO deployment at the edge requires the appropriate hardware. For example, the OnLogic Axial AC101 and Karbon 804 combine Intel's lineage of market leading processors with OnLogic's industrial and embedded computer design to bring scalable, high performance processing to the enterprise network edge. Both systems can be configured with a range of Intel 13th Generation Hybrid-Core processors (formerly known as Raptor Lake).

These multi-core CPUs support Intel Deep Learning Boost (DL Boost), are equipped with integrated UHD 770 graphics, and are capable of supporting up to 128GB of DDR5 memory, which provides native robust AI capabilities when paired with Intel OpenVINO.



Additionally, both systems support the Intel ARC series of discrete GPUs to further elevate the AI processing capabilities. The OnLogic Axial AC101 will also support Intel 14th generation processing and the Intel Flex Series of datacenter-class GPUs.





OpenVINO can assist with easily deploying AI workloads across your various hardware and accelerators depending on your specific project needs.

Conclusion

Intel OpenVINO empowers you to harness the true potential of AI at the edge. By reducing latency, enhancing security, and improving scalability, OpenVINO unlocks a world of possibilities for faster decision-making, improved operational efficiency, and the development of innovative solutions that take AI from potential to practice.

Combined with OpenVINO, the OnLogic Axial AC101 Edge Server or Karbon 804 Rugged Computer provide the ideal platform for on-premise AI inference deployment at the edge. These systems support Intel 13th or 14th generation Core™ processors, with integrated Intel UHD770 GPU, and can be configured with discrete graphics cards, including Intel ARC GPUs, for advanced AI training and inference workloads.

Ready to get started?

Connect with us to explore how running OpenVINO on OnLogic's Edge Ai Computers can transform your business.

North America

Call: +1 (802) 861 2300
Email: info@onlogic.com
www.onlogic.com

Europe

Call: +31 88 5200 700
Email: info.eu@onlogic.com
www.onlogic.com/eu-en/