Developing for the Edge

Edge developers are in demand

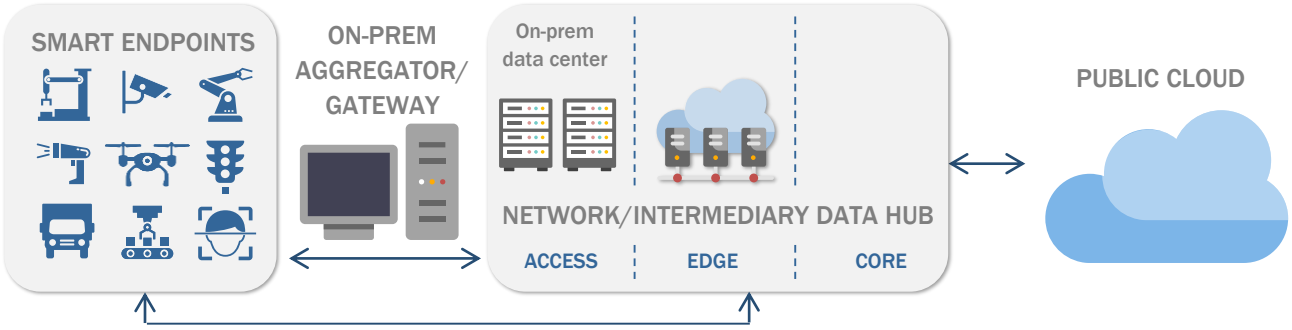
Businesses looking to harness the power of data to maximize efficiency, improve customer experience, increase overall revenue, and improve worker safety, need applications to collect, process, transmit and analyze novel data sources. This enables businesses to process data and make decisions locally, often while maintaining a presence in the cloud for scalable processing power. With 75% of all data projected to be created outside of central data centers by 2025[1], better business outcomes depend on putting compute closer to where data is generated - at the edge. Nearly 73% of Business Decision Makers, Developers, and IT Decision Makers that are aware of edge computing see edge computing as a core or secondary component of their IoT solution.[2] Many of these applications require near-real-time latency in areas where cloud connectivity is unreliable and expensive.

While businesses may have access to the hardware needed to drive computing at the edge, the applications simply aren't there yet. Developers capable of working at the edge are in demand. Those that understand both edge and cloud development are even more valuable as they better understand how their work integrates into the overall ecosystem. Almost half of companies say IoT helps them increase the connectivity of old and new devices (49%) because it aids in protocol translation. Nearly as many report that operations are improved because there is no longer a reliance on internet bandwidth (47%).[2]

This paper explores some of the considerations, challenges and strategies of edge development and presents some solutions to streamline and accelerate edge development.

What exactly is “the edge”?

Generally, edge refers to processing data as close as possible to the point of collection to drive quicker, more efficient insight and action. Often referred to as IoT or connected devices, the idea of edge compute is not new. However, "edge" is not universally defined, and different organizations will have different interpretations. Understanding what one means by “the edge” has dramatic implications depending how far out on the edge code is deployed.



EDGE DEVICES

Also known as connected devices or embedded compute, edge devices (minimally) process data close to the collection point. Often low in power and compute resources, edge devices require that code is deployed close to or on bare metal and control functions. Thus, code needs to be incredibly secure, elegant, and efficient to reduce latency and preserve resources (power, memory) yet anticipatory enough to be self-sufficient.

LOCAL COMPUTE/GATEWAYS

Local compute/gateways operate in a middle ground and provide the necessary power at proximity to enable edge applications. Spanning in capacity and functionality from a PC to a small data center, these local compute options provide some of the cloud comforts: an OS, ability to host containers, and coding in popular/common languages like Java and Python. Local compute also allows for data reduction and preliminary analysis before sending edge data to the cloud for analytics, thus saving substantial time, money, and bandwidth.

ON-PREM DATA CENTERS

An on-prem data center may be considered "edge" by tech companies to differentiate from the public cloud. However, the developer experience is often closer to “cloud” than “edge”. Apart from the ability to leverage some cloud-native services, an on-prem data can look very much like public cloud development. On-prem (or cloud) data centers are necessary for integrating data streams across locations, driving AI and/or analytics models, long term storage and archiving (leveraging cheap cloud storage).

Why develop for the edge?

Edge developers solve for a variety of challenges that can't be addressed solely by cloud.

EXPENSE/INEFFICIENCY OF TRANSMITTING TO THE CLOUD

Organizations can't send all their edge data to the cloud. Sending all data to the cloud can be expensive and inefficient. Much of the data on edge devices are most relevant when used in near-real-time to inform operations; a subset of aggregated data that is valuable on a broader level gets passed to the cloud for analytics/model training.

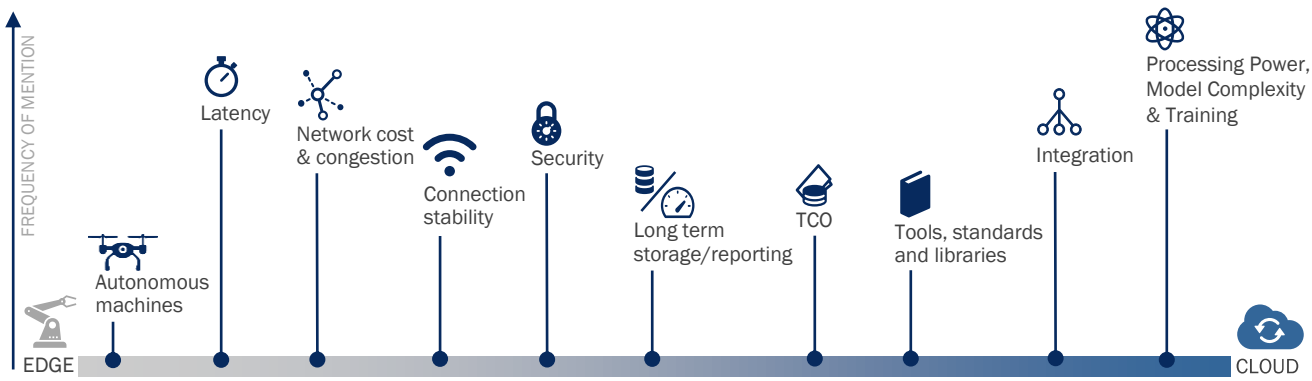
CRITICAL LATENCY ISSUES

Latency is critical and sub-second delays can be catastrophic. When on-site equipment (for example, oil wells, anti-lock braking technology) sends an alert that requires immediate attention, local compute can save companies millions of dollars, particularly for safety and preventive maintenance.

UNRELIABLE CONNECTIVITY

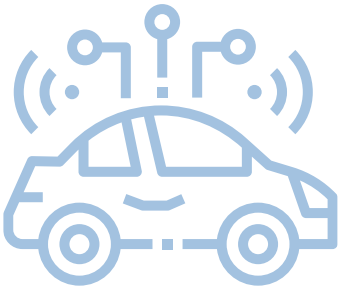
Remote devices may lack reliable internet connectivity. In industries with remote/mobile operations, compute needs to take place locally and push to cloud when the device can get online.

A need for greater integration and processing power require some applications remain anchored in the cloud.



How have developers solved for IoT edge solutions?

IoT edge devices have specific functions and are required to perform these functions in near real-time. Any delay in sending network traffic to another node on the network could result in serious consequences. When developing applications for these devices, software developers can develop it themselves, utilize the open source community for common libraries and code, or take advantage of validated reference implementations. Each use case link shows documentation, including architectural diagrams and recommended development kits.



Fleet Driver Management

Designed to detect and track driver behavior and driver actions to ensure safe driving. This application detects driver's drowsiness states and distractions like phone use. Provides real time alerts to the driver and analytics per driver. Deep learning models, camera video streams, and analytics run on In-Vehicle-Computers.

https://software.intel.com/content/www/us/en/develop/articles/driver-management.html

Edge Insights for Vision

Edge Insights for Vision features a set of pre-integrated ingredients designed for computer vision and deep-learning inference for applications at the edge, optimized for Intel® architecture. Implement as a containerized architecture or a stand-alone runtime.

https://software.intel.com/content/www/us/en/develop/documentation/edge-insights-vision-doc/top.html

Automated Checkout

A common open middleware framework, EdgeX Foundry*, enables multiple IoT sensors to work together and deliver an automated checkout solution that can ultimately scale for an entire store.

https://software.intel.com/content/www/us/en/develop/articles/automated-checkout.html

Enabling intelligence at the edge is achieved through retrofitting existing systems or through new modular infrastructure implementation.

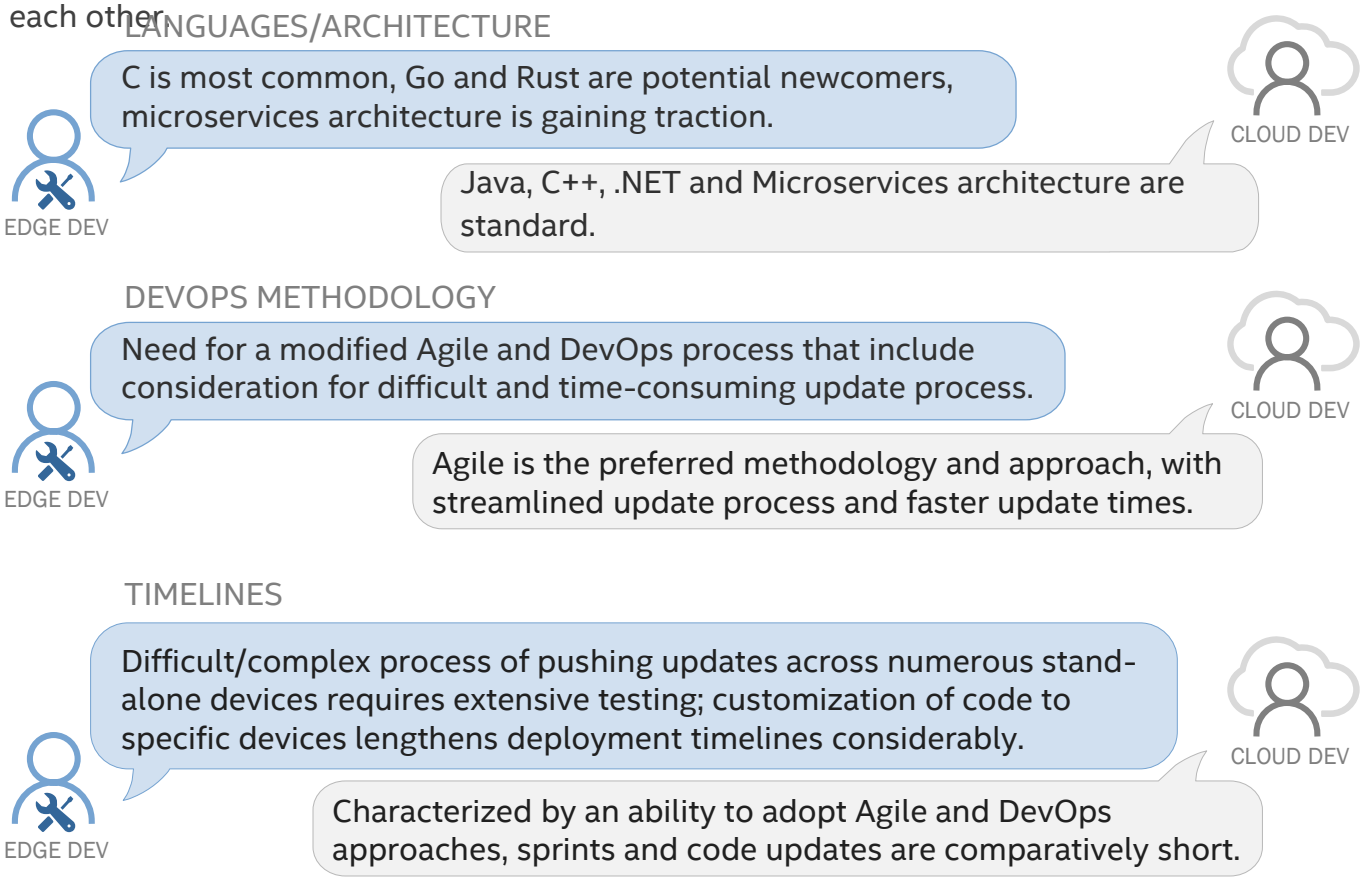
For industries with a long asset life cycle (e.g. manufacturing and transportation), edge intelligence is often enabled via a separate device or board in the device. For example, a separate board in the cabinet, or a data collection device mounted on or nearby equipment solves for many IT-OT convergence challenges:

- New functions don't interfere with primary core functionality.
- Separate swim lanes for IT and OT are preserved.
- Equipment ROI is increased by extending its life.

Similarly, in greenfield situations, new edge deployment is often done modularly, enabling organizations to make long-term investments, allow for scale, and leverage cloud infrastructure.

What does it mean to develop for different edge destinations?

Edge and cloud development can feel like alternate sides of the same coin, often talking past each other.



Developing for different edge destinations *(continued)*

RESOURCE UTILIZATION

Whether writing to the metal or leveraging virtualization and containers, code needs to be power efficient—turning off memory, shutting down processes when not in use, etc.; edge devices hold only the logic they need to do their jobs and scale to cloud as needed.

EDGE DEV

CLOUD DEV

Leveraging power of near limitless compute and storage, developers can focus on business logic without getting bogged down in minutia; cloud apps can hold extraneous "just in case" features/functionality.

DEPLOYMENT/MANAGEMENT

Traditionally, software has been self-contained and manually deployed/ updated. Containers and microservices architecture haven't historically been possible on edge devices but are gaining traction as devices get more powerful/secure and orchestration is enabled.

EDGE DEV

CLOUD DEV

Leveraging containers and microservices architecture allows applications to consume services on an as needed basis. Easy deployment means faster development cycles and regular updates.

TECHNOLOGY OWNERSHIP

Edge applications often straddle IT/ OT and can even feature separation of ownership of software services and physical devices (e.g. airport kiosk are owned by airports, but software is managed by airlines).

EDGE DEV

CLOUD DEV

Cloud applications sit squarely in the purview of IT application development teams.

CHALLENGES

Heterogeneity of devices and environments/ connectivity, resource utilization and remote locations makes it difficult.

EDGE DEV

CLOUD DEV

Staying on top of new services and taking advantage of cloud native capabilities can be hard. Cost control in an unlimited environment is also a challenge.

Companies leverage local compute or gateways to harness and combine edge and cloud benefits

More compute/power/storage resources mean more sophisticated intelligence and ability to code in higher order languages and build on scalable micro service architecture:

Data are reduced/transformed to preserve often limited bandwidth for mission critical data and applications

Updates/security are deployed more easily on fewer, more homogenous devices, managed via containers.

Better models and more complete business intelligence is achieved by pooling data across devices rather than creating individualized models.

Local compute preserves the near-real-time latency requirements by remaining physically proximal to the edge device.

While local compute has more power than edge devices, model training still requires the resources of cloud/data center compute.

For example, Yumei Die Casting Co., Ltd., incorporated defect detection as parts were cast, rather than waiting until the end of the line. They accomplished this via computer vision on one line coupled with machine integration and model training via industrial grade gateways and industrial PCs, yielding dramatic cost savings. Read more.

What are the potential challenges of edge applications?

1. DEVICE HETEROGENEITY & LONGEVITY

Expensive equipment with long refresh cycles leads to device heterogeneity. To maximize the ROI of new functionality and existing equipment, applications need to work with multiple generations of equipment from different OEMs.

2. UNIQUE, HIGH-RISK APPLICATIONS

Edge applications are specific to each organization's situation. Failure at the edge has potential for catastrophic consequences --both financial and safety oriented.

3. OT/IT CONVERGENCE AND SECURITY CONCERNS

Edge equipment is often the realm of OT, while business intelligence, database, security and the network itself are IT. Edge applications bring the two together and introduce security concerns by bringing new equipment onto the network, thereby exposing both the equipment and the network to new risks.

How do we solve for these challenges?

Enterprises often lack a playbook for how to use edge intelligence to transform business models/ processes, and the specific expertise needed to modify existing equipment. Technical decision makers and developers have a critical role to play in solving and innovating for these and ultimately bring edge intelligence to life.

Those advocating for the edge need to be prepared with details on what the solution is (hardware, software, services), how it will integrate with their heterogeneous environment, and how it (and the network it connects to) will be secured.

Edge applications require the right software architecture that combines flexibility, efficiency, and security.

Successful edge deployments require both adoption of technology and organizational change. This means including both IT and OT in the conversation as well as articulating the transformational business benefits of edge intelligence.

Intel has the playbook to help organizations successfully deploy to the edge

The Intel® Edge Software Hub can streamline development and accelerate deployment of innovative solutions for the edge with a one-stop resource for software and optimized use case-specific packages. To learn more, please visit: https://intel.com/edgesoftwarehub.

At the Edge Software Hub, developers can browse, learn, and download pre-packaged software that comes with reference implementation to help enable intelligence at the edge via retrofitting existing systems or new modular infrastructure implementation. Software developers: Learn how you can customize, validate, and deploy solutions faster and with greater confidence by visiting https://intel.com/edgesoftwarehub/develop.

Additional Use Cases and Example Packages

Edge Insights for Industrial

Intel® Edge Insights for Industrial eases AI deployment in industrial environment through a set of integrated key capabilities, such as data ingestion, processing and transmission. The package optimizes edge analytics, driving improvements in industrial product quality, operational performance, prediction of downtime, and automated operational flows. Using Edge Insights for Industrial, developers can extract data from machines, gateways, and devices like cameras at the edge. It also helps communicate information interchangeably and securely, provide cohesive management, and analyzes data quickly in an industrial environment.

LEARN MORE: https://www.intel.com/content/www/us/en/internet-of-things/industrial-iot/edge-insights-industrial.html

Converged Edge Insights and Wireless ready Intelligent Traffic Management for Smart City

This modular, container-first architecture package combines computer vision inferencing and wireless networking workloads on a single device that identifies and manages dangerous traffic intersections. It utilizes Intel® Open Network Edge Services Software toolkit to host a 5G radio access network (RAN) on the same edge device, while enabling edge to cloud workflow. This builds and manages the infrastructure to deploy, monitor, and orchestrate virtualized applications across multiple edge devices with Intel® OpenNESS-Ready Marketplace controller/node architecture.

LEARN MORE: https://software.intel.com/content/www/us/en/develop/topics/iot/edge-solutions/converged-insights-recipes.html

1. What Edge Computing Means for Infrastructure and Operations Leaders, Gartner, Oct 3, 2018
2. Microsoft IoT Signals Report, 2nd Edition, October 2020