Chapter 4
Towards Intelligent Ship–Edge Computing Enabling Automated Configuration of Ship Models and Adaptive Self–Learning

Fearghal O'Donncha
https://orcid.org/0000-0002-0275-1591
IBM Research, Ireland

John D. Sheehan
IBM Research, Ireland

Maroun Touma
IBM, USA

Sofiane Zemouri
IBM Research, Ireland

Rob H. High
IBM, USA

ABSTRACT

Edge computing is a solution that prioritizes data processing for low-latency and specialised applications, particularly those involving complex and mission-critical processes. An ideal example of an edge device is a ship, which heavily relies on its onboard computing capabilities to autonomously perform various navigation tasks. The transportation sector, including shipping, has witnessed the emergence of software-defined vehicles where sensor data, analytics, and algorithms play a pivotal role in optimising operations such as propulsion, cargo handling, energy management, communication, and human-machine interactions. This chapter delves into the significance of edge computing in the shipping industry, outlining its ability to enhance operational efficiencies. It explores the specific user requirements for an effective edge computing solution and highlights the role of AI in enabling scalable and continuous computation. Additionally, the chapter provides a comprehensive overview of edge infrastructure, platform requirements, and considerations pertaining to data and AI at the edge. The discussion incorporates the Mayflower autonomous ship as a case study, illustrating the criticality of edge computing for the future of shipping.

DOI: 10.4018/978-1-6684-9848-4.ch004

INTRODUCTION

Edge computing is an architecture rather than a specific technology. It provides a framework for distributed computing that aims to bring enterprise applications closer to data sources such as local edge servers or IoT devices. By being in close proximity to data sources, edge computing provides several benefits, including faster insights, improved response times, and improved bandwidth availability. It is a response to the increasing demand for large-scale data processing at the edge, and the need for resilience to network disruption.

While this data-centric approach offers many mission-critical benefits, it is crucial for edge computing to exist as part of a cloud continuum that allows stakeholders to interact with data and processes it in a unified way. This ensures that edge computing is integrated seamlessly into existing cloud infrastructure, enabling a unified and cohesive approach to data processing and management.

EDGE COMPUTING OVERVIEW

Edge computing minimises latency by bringing computational workloads closer to the data source and the location where actions need to be taken – removing the distance between collecting data, processing it, and making those results available for use in decision making. Data is generated from diverse sources and at different scales, encompassing equipment, devices, individuals, and processes.

The concept of edge computing is closely intertwined with cloud-native development and decisions regarding where computational tasks should be executed. It prompts considerations on which data should be transmitted back to the cloud for further processing.

Telecommunication providers often refer to the network edge, which offers a chance to deploy edge infrastructure and leverage communication capabilities, particularly with technologies like 5G. This presents an opportunity to enhance edge capacity and optimize the utilization of communication resources.

The integration of compute capabilities into modern industrial equipment has opened up new possibilities for performing a wide range of tasks. For instance, cameras on assembly machines can be utilized for quality control analytics. Production processes can be optimised, comprehensive management of distribution centres becomes feasible, and software containers can be deployed in a suitable runtime environment enabling elastic scalability.

However, managing these deployments poses significant challenges. With an estimated 15 billion devices worldwide, enterprises must oversee the provisioning, deployment and ongoing management of thousands of devices, each with different specifications, purposes, operating systems, and workloads. While the provisioning process can be done in a controlled environment to ensure full compliance, security becomes a major concern as these devices operate outside the traditional confines of IT data centres. They lack the typical safeguards found in hybrid cloud environments, such as physical barriers and the uniformity required for certification. Ensuring the integrity and preventing tampering of edge devices becomes paramount.

To support the growing role of edge computing, it is crucial to build an ecosystem that facilitates its development. The impact of edge computing on enterprise computing is expected to be substantial, comparable to how mobile technology has transformed the consumer sector.

Cloud vs. Edge Computing

The primary difference between cloud computing and edge computing is the location where data processing occurs. In cloud computing, data is processed on a central cloud server, which is usually located far away from the source of information. Traditionally compute took place on centralized cloud services such as Amazon EC2 instances. Hybrid cloud is a mixed computing environment where applications are run using a combination of computing, storage, and services in different environments—public clouds and private clouds, including on-premises data centres or edge locations.

Hybrid cloud computing approaches are widespread because almost no one today relies entirely on a single public cloud. Hybrid cloud solutions offer the flexibility to seamlessly migrate and manage workloads across diverse cloud environments, empowering organizations to tailor their infrastructure to meet specific business requirements. By adopting hybrid cloud platforms, organizations gain the ability to lower costs, mitigate risks, avoid vendor lock-in, and leverage existing cloud-native developer skills and CI/CD pipelines to drive successful digital transformation initiatives.

In today's landscape, the hybrid cloud approach has become a prevalent infrastructure setup. As organizations undergo cloud migrations, hybrid cloud implementations naturally emerge, enabling a gradual and systematic transition of applications and data. With hybrid cloud environments, enterprises can continue utilizing on-premises services while harnessing the advantages of public cloud providers like AWS, Azure, and GCP, which offer flexible options for data storage and application access (Google, n.d.).

Edge computing extends the hybrid cloud paradigm to address the unique requirements of enterprise and consumer applications. Although they possess individual traits, edge computing and hybrid cloud can collaborate to establish a comprehensive and adaptable computing infrastructure. A notable aspect of contemporary edge computing solutions is their adoption of cloud native development practices specifically designed for the edge. By leveraging cloud native development practices, applications can be built and deployed using the same skills and tools that have been honed for developing cloud native applications in hyper-scale cloud environments or private data centres. This allows for the seamless extension of these practices to the edge, enabling organizations to leverage their existing expertise and resources for edge computing deployments.

Hence, aspects such as distributed computing, data processing, management, and integration, as well as workload placement and optimisation are enabled in an accelerated and scalable manner.

Concept and Technology

Edge architecture encompasses all the active components of edge computing, including devices, sensors, servers, and clouds, spread throughout the network wherever data is processed or utilized. By bringing processing closer to the data source, edge architecture significantly reduces latency and cost: applications and programs running at the edge can swiftly and efficiently respond to user interaction and data without the need for transferring data across a wide area network, resulting in an enhanced user experience and optimal overall performance.

The definition of the "edge" is flexible and context-dependent. For example, in the case of a shipping company, the edge may encompass activities at docks where shipments are loaded and unloaded, or it may include on-board processing of ship systems and operations. In both scenarios, processing and analysis occur in near real-time, leading to data-driven decision-making. While the company's headquarters with the main data centre may be located miles away, the edge represents the crucial point where data is collected, processed, and managed to derive insights, irrespective of latency challenges.

Why Edge Computing in the Maritime Industry

Ship transport accounts for over 80% of global trade in goods and raw materials, making the efficiency and speed of large container ships critical to the value chains and production processes of countries worldwide. However, the environmental impact of shipping, including emissions, fuel efficiency, and noise pollution, is becoming more apparent. To address these concerns, emerging technologies like artificial intelligence (AI) and digital twin are being leveraged to improve the efficiency, management, and environmental sustainability of shipping routes.

Whilst digitalisation of the transport sector is key to unlocking efficiencies, this is currently hindered by data silos, technological limitations, and barriers of complexity. DT4GS promises to enable stakeholders in shipping to actively embrace the full spectrum of Digital Twins innovations to support smart green shipping in both the upgrade of existing ships, as well as the building of new vessels.

These solutions will be deployed across the cloud continuum from centralized public cloud services to decentralized edge computing devices. At one end, ship owners and stakeholders can leverage public cloud services provided by major cloud providers that offer centralized computing power and storage resources that can be accessed from anywhere. On the other end, edge computing will be deployed closer to where the data is generated to provide compute at every step of the ship system. This provides unique benefits and advantages in terms of data sovereignty and security, resilience to outages and connectivity constraints, innovation opportunities and flexibility to different providers, as well as inherent autonomy that are critically important to shipping.

There are multiple complexities inherent to ship operations and decision making. Whilst the central complexity may be the difficulty and uncertainty inherent in decision making in the ocean, there are multiple technological challenges to overcome:

- Scale: There are thousands of ships with limited or no technical expertise on board. The solutions implemented must be self-healing and turn-key, requiring minimal intervention and maintenance.
- Heterogeneity: Each ship represents a dynamic and unique environment, making it challenging to find a one-size-fits-all model. Solutions need to be adaptable and flexible to cater to the diverse needs and characteristics of each vessel.
- Data Gravity: Ships and their instruments generate vast amounts of logs, events, and metric data, often in different formats. Additionally, the sensitivity of some of this data adds another layer of complexity. Handling and processing this data in a secure and efficient manner is crucial.
- Resource Constraints: Ships serve as the ultimate edge nodes, characterized by limited communication, computational, and energy resources. These constraints must be considered when designing solutions, ensuring they can operate effectively within the ship's resource limitations.

Fundamentally, edge and cloud technologies promise a solution that seamlessly infuse AI and simulation across the entire spectrum of shipping operations. Agnostic of connectivity or compute capabilities, it provides a continuously available digital assistant to guide all aspects of decision and automation.

AI can assist Captains and crew to make better decisions by seeing things humans overlook or don't fully recognize; by assessing alternatives they may not have considered; by evaluating and optimizing choices with more mathematical precision than can be accomplished by the human mind alone; by reacting to events in real-time and at a rate that is much higher than is humanly possible. Perhaps the AI can pre-empt these crises from emerging. And the holy grail of these potential advantages is to inspire decision-makers to ideas they might not have produced on their own.

Doing so promises to improve decision-making, which in turn will improve operational efficiency, safety, and decrease burden. With AI we can reduce the current carbon footprint of transporting goods across the oceans. We can incur fewer accidents, limit cargo damage and loss, and most importantly make maritime travel less risky.

To do this, we need to harness the power of AI to recognize dangers, evaluate alternatives, and present recommendations in the decision-making process. We need to incorporate AI into the human decision processes. In essence, the AI needs to act as another member of the crew, offering interpretation, advice, and warnings as appropriate to provide meaningful contribution to how crew leaders and the Captain make their decisions. Easier said than done.

A prominent example in this regard is the Mayflower Autonomous Ship (MAS400) (Kıcıman et al., 2023; MAS, n.d.). The MAS400 vessel was built without any provision for humans on board and must make all navigational decisions on its own. It uses an array of visual and other signal input to sense its environment. AI is used to perceive and interpret what it "sees". It classifies a wide range of objects and determines whether they are an obstacle to its current course. It establishes potential strategies for navigating around them, and evaluates each for their safety, conformance to regulations, and optimization. And then, of course, it puts those decisions into action – controlling rudder and speed to achieve the outcome it has determined.

However, a potentially larger and more immediate use of this technology is not for vessels that drive themselves all over the world, but rather to assist sea Captains in their own navigational and ship operations responsibilities. One of the lessons from the Mayflower project is that sometimes you do need a human on-board – to fix a broken manifold coupler to the auxiliary engine, or to repair the electrical connection that has frayed and causing a short. And with human lives at risk, you need humans that have the authority and responsibility for protecting those lives – for making the critical decisions necessary to fulfil their mission safely.

STATE OF THE ART IN EDGE COMPUTING

A forecast by International Data Corporation (IDC) estimates that there will be 41.6 billion IoT devices in 2025, capable of generating 79.4 zettabytes (ZB) of data (Hojlo, 2021). The latency advantages of edge computing provides obvious appeal to extract insight from these vast volumes of data. However there are also multiple other considerations driving edge adoption including security, operational costs, resilience, and flexibility (Rathore et al., 2022).

The global edge computing industry is experiencing rapid growth, driven by advancements in technologies such as 5G, IoT, and the Industrial Internet (Rathore et al., 2022). Edge computing entails extending a consistent computing environment from the core datacentre to physical locations in close proximity to users and data. Similar to a hybrid cloud strategy that enables organizations to run workloads across their own datacentres and public cloud infrastructure, an edge strategy expands the reach of a cloud environment to numerous additional locations. This approach allows for a more widespread and distributed computing infrastructure, bringing computing capabilities closer to the edge for enhanced performance and responsiveness.

There are three categories of edge use cases (Schabell, 2022):

- The first is called enterprise edge, and it allows customers to extend application services to remote locations. It has a core enterprise data store located in a datacentre or as a cloud resource.
- The second is operations edge, which focuses on analysing inputs in real time (from Internet of Things sensors, for example) to provide immediate decisions that result in actions. For performance reasons, this generally happens onsite. This kind of edge is a place to gather, process, and act on data.
- The third category is provider edge, which manages a network for others, as in the case of telecommunications service providers. This type of edge focuses on creating a reliable, low-latency network with computing environments close to mobile and fixed users.
- For industry applications, edge computing capabilities can be decomposed into two categories:
- **Edge Server**: This typically refers to IT equipment specifically designed for computing tasks at the edge and operating as an extension to the computation tasks performed in the cloud. It can take the form of a half rack comprising 4 or 8 blades, or it can be an industrial PC. Essentially, it is a piece of IT equipment dedicated to edge computing operations.
- **Edge Device**: An edge device is a piece of equipment built for a specific purpose, such as an assembly machine, a turbine, or a car. While these devices were initially designed for their primary functions, they also possess computing capacity. In fact, many devices that were traditionally considered IoT devices now incorporate increased computational power. For instance, an average car today is equipped with approximately 50 CPUs. These devices often run on the Linux operating system, allowing for the deployment of software containers, and enabling edge computing applications to be executed directly on the devices themselves.

Deploying workloads across both cloud and edge resources presents several fundamental challenges for technology solutions:

- Consider the scale of the challenge at hand. With an estimated 15 billion devices in the world today, enterprises are faced with the task of managing thousands of devices within their operations.
- Heterogeneity at the edge poses another significant factor to address. These devices come in various forms, each serving a different purpose and running on different operating systems. Managing the diverse range of devices and the specific tasks they perform can be a complex endeavour.
- Security is a critical concern when dealing with edge devices. Unlike traditional IT data centres, these devices operate outside the confines of controlled environments. They lack the physical barriers, uniformity, and consistency typically found in hybrid cloud environments that aid in ensuring security. Protecting edge devices from tampering and unauthorized access becomes a priority.
- Building a robust ecosystem is essential. The role of edge computing is rapidly expanding and will have a profound impact on enterprise computing, much like how mobile technology has influenced consumer behaviour. Establishing a comprehensive ecosystem that integrates and supports edge devices and their functionalities will be crucial for organizations to maximize the benefits and potential of edge computing.

- Maintenance and ongoing support presents some unique challenges. Software and AI components are complex and require period patches and updates. Software performance can also be impacted by environmental conditions. Having a rigorous approach for software life-cycle management and drift detection is essential to achieve optimal performance and avoid catastrophic failure where nearby IT support personnel is not available.
- Ensuring data privacy throughout the deployment process is essential to safeguard sensitive information against data breaches, regulatory compliance, and to enhance public trust within organisations. As workloads are distributed across diverse and interconnected environments, ranging from public clouds to edge devices, the potential attack surface widens significantly. Adhering to stringent data privacy measures becomes imperative to mitigate risks associated with unauthorized access, data leakage, and malicious activities.
- The strategic placement and optimization of workloads play a pivotal role in maximizing performance, minimizing latency, and optimizing resource utilization. The deployment of workloads across diverse environments, encompassing cloud data centres and edge devices, presents unique challenges and opportunities. By harnessing the power of intelligent workload orchestration, organizations can unlock the full potential of their computing infrastructure, ensuring optimal performance and resource allocation while seamlessly catering to the unique requirements of individual workloads.

FULL STACK EDGE SOLUTION

Edge solutions require multifaceted capabilities, including the orchestration of heterogeneous workloads, creation of a scalable, extensible, and robust data management pipeline, and seamless deployment of data inference and AI on the edge. The latter consideration is especially critical as many organisations become AI-driven enterprise and infuse generative AI capabilities across their operations (Yusuf, 2023). The requirements of workload orchestration, data operations, and AI infusion, include:

- **Edge orchestration:** A fundamental requirement is to enable seamless deployment of containerised applications to multicloud and edge environments. Deploying on 1000s of nodes in a scalable, replicable manner requires deep automation for advanced management of 1000s of different software applications across 1000s of compute nodes distributed across the entire footprint of an organisation's operating remit. Solutions such as Open Horizon (Open Horizon, 2023) enables the autonomous management of more than 10,000 edge devices simultaneously. Core architecture considerations include:
    - The provisioning of Open Horizon includes the management hub that runs in an instance of OpenShift Container Platform installed in the data centre. The management hub is where the management of all remote edge nodes (edge devices and edge clusters) occurs.
    - Managed edge nodes can then be installed in remote on-premises locations to make application workloads local to where critical business operations physically occur, such as docks, ships, factories, warehouses, retail outlets, distribution centres, and more.
- **Data Operations (DataOps):** Edge consist of disparate devices such as engines, propellors, hulls, berths, shipping containers, etc., that are all generating data. Resilient edge solutions require robust data management and organisational strategies for collecting and handling data, ensuring

compliance with data sovereignty regulations, and providing flexible data quality solutions facilitating the training and deployment of AI models. Further, cognisant of data latency restrictions, data processing on the edge must translate from vast volumes of raw data to digestible subsets of high-quality, high-value features that can be used for AI model finetuning and prediction. As an example, computer-vision based hull monitoring solutions generate large volumes of data, while only small subsets of this data may be relevant for guiding hull cleaning and maintenance.

- **AI Operations (AIOps):** AI-backed decision making will play a critical role in improving the sustainability, safety, and efficiency of shipping. Ship captains and stakeholders can use AI to process large data volumes and make decisions related to route selection, navigation, port arrival and logistics, weather-enforced disruption and mitigation. Further, many aspects can be fully automated such as power management, HVAC system optimisation, cargo and inventory management, predictive maintenance, crew management, and risk analytics. As AI begins to play a central role in shipping, it is critical that 1) ships possess the compute infrastructure to fully exploit these advantages and 2) the AI layer provides reliable inference monitoring mechanisms to assess model performance, detect potential errors or uncertainties, and instil confidence in the decision-making process.
- **Network Operations (NetOps)** plays a crucial role in the seamless deployment of workloads across the cloud to edge continuum. NetOps makes sure the network infrastructure is properly configured, optimized, and secured. It involves activities like network surveillance, traffic control, performance enhancement, security setup, and troubleshooting. Ship owners may ensure effective data transmission, low latency, and dependable connectivity across cloud and edge settings by utilizing NetOps principles, enabling the smooth deployment and execution of workloads. Additionally, NetOps aids in resolving issues with network resilience, scalability, and resource allocation, enabling businesses to fully utilize the cloud to edge continuum. Ultimately, NetOps serves as a crucial link that makes it possible for workloads to be seamlessly integrated and run across a range of computer platforms, improving flexibility, performance.

Edge Infrastructure and Compute

Recent years has seen the evolution and maturation of IoT technology from primarily a data collection and transmission approach, to devices with significant compute capabilities and the ability to deploy workloads to the data. In traditional IoT setups, data generated by connected devices were often transmitted to the cloud for processing and analysis. This approach worked well for certain applications but introduced challenges like latency, network congestion, and increased reliance on cloud connectivity. Edge computing emerged as a solution by bringing data processing closer to the source, enabling real-time analysis and quicker response times.

As edge matures, a critical concern for stakeholders is infrastructure considerations such as server selection for different workloads, security of distributed devices, and loading necessary software onto hardware devices to support various edge applications such as visual inspection, voice interaction, and inference.

Compute capabilities within the shipping and logistics space is an evolving state. While container ships may have some computing equipment for basic tasks such as monitoring and controlling ship systems, these capabilities are generally limited in scope and processing power. The computing infrastructure on a container ship is designed to support essential functions like engine control, navigation, communication, and safety systems.

However, it's worth noting that with the increasing adoption of digital technologies and automation in the shipping industry, there is a growing trend towards incorporating more advanced computing capabilities on certain types of vessels. For example, larger and more advanced container ships may have additional computing systems for tasks like cargo management, route optimization, and fuel efficiency monitoring.

Overall, while there may be some level of compute capability on a typical container ship, it is typically limited and focused on specific operational requirements rather than extensive computing tasks. The bulk of compute-intensive tasks, data processing, and analytics are more commonly performed in onshore data centres or edge computing systems that support the shipping operations.

Several companies offer edge solutions that provide integrated hardware and software for edge computing deployments. Some examples include:

- Dell EMC Edge Gateway provides a compact and ruggedized device designed for edge computing. It comes with off-the-shelf, pre-configured, pre-certified, ready-to-use products, for different industries or applications (Dell, 2023).
- Intel NUC Edge is a compact edge device that promises an out-of-the-box solution ideal for running any critical applications on-premise with immediate high availability (Scale Computing, n.d.).
- Microsoft Azure Stack Edge is a solution that combines hardware and software to bring AI, analytics, and IoT capabilities to the edge. It includes an on-premises appliance that can be deployed in disconnected or low-connectivity environments and integrates with Azure services for seamless cloud-edge integration (Microsoft Azure, n.d.).

When one considers compute infrastructure for shipping, it is important to consider the types of compute capabilities that may be deployed: edge servers and edge devices (described in more detail in Section 1).

Edge Platform and Orchestration

Edge, by its very nature, generally involves 1000s of devices. Hence, concepts from the data centre do not translate directly to the edge. Instead, it requires software capabilities to monitor and update a limitless number of edge devices from across the world and new security technology and protocols to keep everything safe (Hines, 2023).

An edge computing platform contains many components. These include:

- **Edge Management and Orchestration:** This component handles the management, configuration, and physical deployment of edge devices and placement of applications that will run on them. It ensures efficient resource allocation, software updates, and monitoring of edge infrastructure.
- **Security and Authentication:** Edge computing software incorporates security measures to protect edge devices, data, and communications. This includes authentication, encryption, access control, and threat detection capabilities.

- **Edge Gateway:** The edge gateway serves as a bridge between the edge devices and the central cloud or data centre. It provides connectivity, protocol translation, and data aggregation capabilities.
- **Containerization and Virtualization:** To enable portability and ease of deployment, containerisation of software workloads is critical. Software containers or virtual machines encapsulate applications and their dependencies, making them easier to manage and deploy at the edge.

Edge management and orchestration are critical components of edge computing. Edge management refers to the management of edge devices, including device provisioning, configuration, and monitoring. Edge orchestration refers to the management of network resources across an edge network by controlling how network resources flow between devices and applications in an edge environment to produce a more responsive and smartly optimized network.

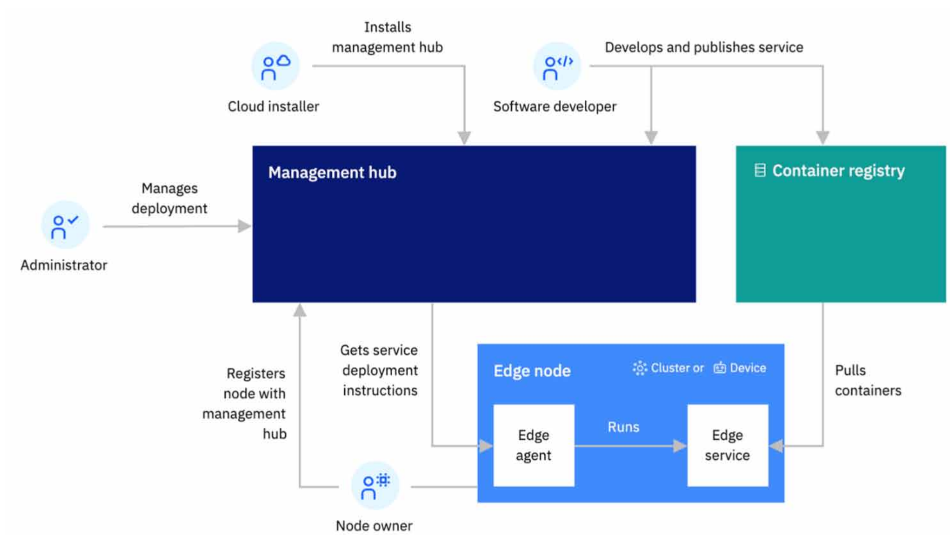
As an example, Open Horizon deploys an orchestrator called "management hub" in an instance of OpenShift Container Platform installed in the data centre. The management hub is responsible for controlling the placement of containers to all remote edge nodes, both edge servers and edge devices.

Figure 1 provides a typical edge computing architecture. The two most critical responsibilities of the orchestrator are the deployment and monitoring of workloads.

Fundamentally, the edge deployment is based on containerisation of workloads. A Cloud Operations team installs the management hub components. Specific applications are developed by data scientists and domain experts and containerised, which are then published to the management hub edge library. Administrators define the deployment policies that control where edge services are deployed.

Edge orchestrators use various monitoring techniques to ensure that edge devices are functioning correctly and that network resources are being used efficiently. Some examples of monitoring techniques used by edge orchestrators include:

Figure 1. High-level topology for a typical edge computing setup
Taken from https://open-horizon.github.io/



- Device monitoring: Monitor the status of edge devices to ensure that they are functioning correctly. This includes monitoring device health, connectivity, and performance.
- Application monitoring: Monitor the performance of applications running at the edge of the network to ensure that they are meeting performance requirements.
- Network monitoring: Monitor network traffic to ensure that network resources are being used efficiently and that there are no bottlenecks in the network. This can also include monitoring of network traffic for security threats and vulnerabilities.
- Data monitoring: Edge orchestrators monitor data flows between devices and applications to ensure that data is being processed correctly and that there are no data integrity issues.

These monitoring techniques help edge orchestrators to identify issues before they become critical and to optimize network resources for better performance. In the context of shipping, it is critical to consider “Day 0/Day 1/Day 2” of the software lifecycle:

- Day 0 activities encompass the initial definition, design, procurement and provisioning of the hardware and software solution. For factory pre-install, this stage includes the initial installation and testing of the hardware and software in the factory prior to shipping it to its final destination. Additional testing are also performed when the device is wired into the network at its final destination and powered up.
- Day 1 marks the actual deployment or launch of the software solution. It represents the first day of production use or operation. On Day 1, the software is made available to users or customers, and the system goes live. This phase involves activities such as testing, data migration, user onboarding, and initial training. The goal of Day 1 is to successfully deploy software created during Day 0 and transition from the development phase to the operational phase and ensure that the software meets the required functionality and performance standards.
- Day 2 refers to the ongoing operational phase of the software after it has been deployed and is in active use. It represents the period of maintenance, support, and continuous improvement. During Day 2, organisations focus on tasks such as monitoring, troubleshooting, bug fixes, performance optimization, regular updates, and feature enhancements. The emphasis is on managing and maintaining the software to ensure its stability, reliability, security, and efficiency throughout its lifecycle.

In the case of shipping, the vast number of deployed edges and the geographical location of each deployment generally makes it prohibitively expensive to provide local IT support at each location to monitor, maintain and update these deployments. Instead, edge technologies must provide a comprehensive solution for administration, managing, monitoring, and securing an almost limitless number of edge servers and devices.

The deployment of thousands of Edge devices means that each of those devices are potential entry points for hackers and security breaches. There are some critical considerations to security on the edge. Some of these include (Iyengar, 2023):

- The enterprise should have the ability to check whether the edge nodes are operating properly by comparing the current configuration of various resources against the desired state.
- The communication between an edge agent and management hub should be signed and encrypted.

- Each container run on an edge endpoint should be verified against the official container to check for tampering.
- Each container should be running in its own "docker" network with self-regulated (application dependencies) connectivity between the components.
- Each edge agent should check that it is running the latest version of the container when the endpoint is connected to a network.
- The enterprise should create a cryptographic signing key pair and have a plan to rotate them.

In the shipping industry, security plays a vital role as any unauthorized access or control over a vessel's ecosystem can have catastrophic consequences. Ensuring robust security measures is of utmost importance to prevent potential threats and protect the safety and integrity of maritime operations.

In contrast to typical centralized Internet of Things (IoT) platforms and cloud-based control systems, the edge control plane is mostly decentralized. Each role within the system has a limited scope of authority so that each role has only the minimum level of authority that is needed to complete associated tasks. No single authority can assert control over the entire system. A user or role cannot gain access to all nodes in the system by compromising any single host or software component. For Open Horizon, the control plane is implemented by three different software entities (IBM, 2021):

- Horizon agents: Outnumber all of the other actors in Open Horizon. An agent runs on each of the managed edge nodes. Each agent has authority to manage only that one, single, edge node. The agent advertises its public key in Horizon exchange, and negotiates with remote agbot processes to manage the local node's software. The agent only expects to receive communications from the agbots that are responsible for deployment patterns within the agent's organization.
- Horizon agbots (agreement robots): Are processes that can run anywhere. By default, the processes run automatically. Agbot instances are the second most common actors in Horizon. Each agbot is responsible for only the deployment patterns that are assigned to that agbot. Deployment patterns consist primarily of policies, and a software service manifest. A single agbot instance can manage multiple deployment patterns for an organization. Deployment patterns are published by developers in the context of an Open Horizon managed user organization. The deployment patterns are served by agbots to Horizon agents. When an edge node is registered with Horizon exchange, a deployment pattern for the organization is assigned to the edge node. The agent on that edge node accepts offers only from agbots that present that specific deployment pattern from that specific organization. The agbot is a delivery vehicle for deployment patterns, but the deployment pattern itself must be acceptable to the policies that are set on the edge node by the edge node owner. The deployment pattern must pass signature validation, or the pattern is not accepted by the agent.
- Horizon exchange: Refers to a centralized, but geographically replicated and load balanced, service that enables the distributed agents and agbots to join and negotiate agreements. Horizon exchange also functions as a shared database of metadata for users, organizations, edge nodes, and all published services, policies, and deployment patterns.
- The switchboard in Open Horizon serves as a central communication hub, facilitating coordination and orchestration between edge devices and the management infrastructure. It ensures efficient workload placement, dynamic updates, and seamless communication, optimizing resource utilization and enabling effective deployment of workloads.

- The MMS/Sync service in Open Horizon is responsible for synchronizing and managing edge devices within the ecosystem. It ensures devices stay updated with the latest configurations, policies, and workloads from the central management hub, enabling efficient and secure communication. By keeping devices in sync, it enables seamless coordination and orchestration of workloads, enhancing the scalability and performance of edge computing deployments.

To maintain anonymity, the agent and agbot processes share only their public keys throughout the entire discovery and negotiation process. All parties within Open Horizon treat each other party as an untrusted entity by default. The parties share information and collaborate only when trust is established, negotiations between the parties are complete, and a formal agreement is established.

Edge-in-a-Box

Edge-in-a-box takes a full stack approach to deploying Machine Learning and AI models on the edge in order to provide a high level of resiliency and semi-autonomous operations. It defines a delivery model and associated tools and capabilities for users who want to run complex analytics on premises, and closer to where the data is generated. Enterprises today face a number of challenges in scaling Edge deployments to numerous locations. Two main hurdles encountered are:

- *Time/cost to deploy*: Each deployment consists of several layers of hardware and software that need to be installed, configured and tested prior to deployment. Today, a service professional can take days or even weeks when installing complex AI solutions *at each location* - severely limiting how fast and cost effectively enterprises can scale up their deployments.
- *Day-2 Management:* The vast number of deployed edges and the geographical location of each deployment could often make it prohibitively expensive to provide local IT support at each location to monitor, maintain and update these deployments.

To address these challenges, Edge-in-a-box uses a hub-and-spoke approach for (a) zero-touch provisioning of each edge server, (b) continuous monitoring of its state and (c) capabilities to manage & push software/security/config updates to numerous edge locations - all from a central location. Automation of the provisioning process is accomplished using configuration policies that are developed and tested in the factory and ensure that all elements of the stack including compute nodes, storage, operating system, applications and all their dependencies work properly together. As such, prior to installation, each managed edge-in-a-box cluster is attached to the hub and is assigned a placement tag. The placement tag determines what policy set is applied to the attached cluster and, in turn, triggers the installation of all the software components that are associated with the configuration policies in the policy set.

DataOps at the Edge

DataOps is an approach that combines practices, processes, and technologies to enhance data analytics by focusing on quality, speed, collaboration, and continuous improvement. It leverages agile methodologies and automation from software engineering to deliver high-quality data quickly and securely. DataOps extends the concept of DevOps, which integrates software development and IT operations, to the realm of data analytics.

In the era of data-driven enterprises, DataOps has gained significance as organizations recognize the strategic value of utilizing their data assets to make informed decisions. The objective of DataOps is to transform data from a mere liability requiring storage and management into a valuable asset that improves decision-making processes. While centralized cloud solutions face challenges in implementing DataOps, data lake solutions aim to address these challenges. In decentralized edge deployments, data management and mining become more complex, requiring comprehensive solutions to maximize the value of data.

Ships generate vast amounts of data related to systems, engineering metrics, sensor logs, employee information, camera and audio data, as well as external weather and supply chain information. Ideally, a scalable data unification policy is desired to merge these diverse data sources across key business entities such as ship operations, components, navigation, crew, cargo, customers, and suppliers. Traditional approaches relied on simple extract, transform, load (ETL) systems and rule-based merging schemas. However, as the data ecosystem grows larger, more complex, and heterogeneous, machine learning-based tools become essential for effective data utilization in enterprises.

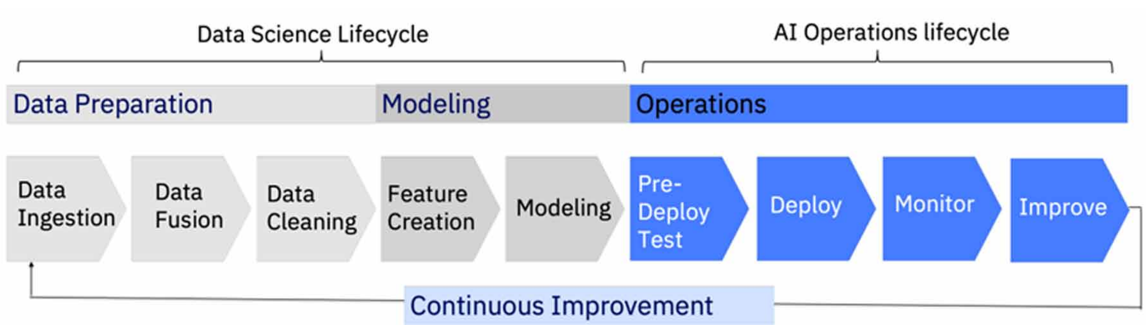
In distributed edge deployments, it is inefficient to transfer all data to the cloud. Instead, data is typically stored locally, processed on edge devices, and only intelligent subsets or insights are communicated to the cloud. DataOps implementation must also comply with relevant regulations and ensure the protection of sensitive data by automating and integrating data governance, data quality, and policy management. In the shipping industry, this consideration extends to adhering to country-specific regulations across different nodes of the system.

AIOps at the Edge

Coined by Gartner, AIOps—i.e. artificial intelligence for IT operations—is the application of artificial intelligence (AI) capabilities, such as natural language processing and machine learning models, to automate and streamline operational workflows. The AI lifecycle consists of multiple stages, including data preparation, modelling, and operations. Figure 2 illustrates a typical model development and deployment pipeline (Arnold, 2020). While significant attention has been devoted to the initial data science stages of the lifecycle, the subsequent stages of AI operations are frequently disregarded or overlooked altogether, despite their essential role in effectively deploying AI models in real-world scenarios.

After deployment, ensuring the generation of reliable predictions by AI models and promptly identifying any erroneous outputs is crucial. This plays a vital role in instilling confidence in the decision-

Figure 2. End-to-end AI lifecycle



making process of the model. In essence, AIOps focuses on the continuous monitoring and improvement of deployed models. AIOps uses big data, analytics, and machine learning capabilities to automate some aspects of the following (IBM, n.d.):

- Collect and aggregate the huge and ever-increasing volumes of data generated by multiple IT infrastructure components, application demands, and performance-monitoring tools, and service ticketing systems
- Intelligently sift ‘signals’ out of the ‘noise’ to identify significant events and patterns related to application performance and availability issues.
- Diagnose root causes and report them to IT and DevOps for rapid response and remediation —or, in some cases, automatically resolve these issues without human intervention.

AIOps encompasses data engineering, machine learning engineering, systems engineering, and reliability engineering, aiming to deliver comprehensive AI solutions tailored to specific use cases. It enables the deployment of AI models in production environments and facilitates rapid updates to adapt to evolving conditions. Initially designed for data centre deployments, AIOps plays a crucial role in edge deployments that face increased exposure, limited protections, and greater isolation from support staff and software updates.

AI on the Edge

In recent years, the field of AI has experienced a tremendous surge in interest and development. It has been described as reaching its “Netscape moment,” indicating a significant breakthrough (Kahn, 2023). This has led to the rapid development and deployment of large AI models by organizations, with a focus on automating and enhancing various aspects of their operations.

Foundation models, also known as pre-trained language models, have emerged as a fundamental component in this AI revolution (Bommasani, 2021). These models serve as the building blocks for a wide range of natural language processing (NLP) tasks. They are trained on extensive amounts of text data from the internet and are specifically designed to comprehend and generate human-like text.

Built on transformer architectures, foundation models excel at capturing contextual relationships between words, enabling them to generate coherent and contextually relevant responses. They serve as the starting point for developing task-specific models, thus earning the term “foundation.” While they have gained recognition for their performance in language-related tasks, they can also be customized for various domains beyond NLP, including computer vision, IT operations, and industry applications. For instance, IBM Research has developed a geospatial foundation model that learns from raw satellite imagery to create customized maps of natural disasters and environmental changes (Raghavan & Shim, 2021). This pre-trained model can be fine-tuned for downstream applications like disaster response, supply chain logistics, and agriculture.

By leveraging foundation models, developers and researchers can save significant time and resources that would otherwise be required to build and train models from scratch. These models provide a robust starting point with generalizable capabilities, which can be further refined and adapted to specific applications or industries.

Prominent examples of foundation models include OpenAI's GPT (Generative Pre-trained Transformer) models, notably GPT-3. These models have garnered attention for their remarkable language generation capabilities and their versatility across a wide range of NLP tasks.

Foundation AI for Shipping

AI can assist Captains and crew to make better decisions by seeing things humans overlook or don't fully recognise. However, it requires the ability to make robust decisions in a complex and chaotic environment. Nearly all contemporary examples of AI at best perform a form of inductive reasoning — recognizing examples of prior knowledge. AI excels at recognizing ships and buoys, rocks and paddle boards, but only because it has been taught what those things look like from training data that exemplifies them. Prior knowledge.

AI struggles to deduce what it is seeing actually means — that when it sees a person on paddle board and then suddenly that person disappears that that means they have fallen and are probably still in the water. It is difficult for AI to extrapolate causal relationships that we all take for granted. We understand that people don't just simply disappear in thin air. We understand that waves can hide things from view. We have an intuition of the laws of physics. We can deduce that when a paddle-boarder suddenly disappears they are likely just low to the water, hidden by waves, and probably haven't drifted far from where they fell.

The renowned computer scientist and philosopher, Judea Pearl emphasizes the significance of causation in the context of AI and deep learning. He argues that causation is essential for the success and advancement of deep learning models because it enables them to move beyond mere pattern recognition and make more accurate and reliable predictions (Pearl, 2009).

According to Pearl, traditional machine learning methods, including deep learning, focus primarily on correlation. They excel at identifying patterns and relationships in data but struggle to determine the cause-effect relationships that underlie those patterns. This limitation restricts their ability to provide explanations or interventions based on a true understanding of how variables interact and influence each other (Pearl, 2009).

To overcome this limitation, Pearl advocates for integrating causal reasoning into deep learning models. By incorporating causal knowledge and understanding the cause-effect relationships in the data, models can go beyond prediction and generate more meaningful and reliable insights. Causal reasoning enables deep learning models to answer "what if" questions, perform counterfactual reasoning, and make interventions in complex systems (Pearl & Mackenzie, 2018).

By embracing causality, deep learning models can achieve a higher level of interpretability, transparency, and robustness. They can provide explanations for their predictions, handle situations where training and test data distributions differ, and generalize better to unseen scenarios. Incorporating causal reasoning into deep learning opens up new possibilities for more reliable decision-making, understanding complex phenomena, and addressing challenges such as bias and fairness in AI systems.

Despite significant advancements in the field of causality over the past two decades, its practical application in real-world scenarios still falls short compared to deep learning-based methods. However, the emergence of foundation models that are pre-trained on extensive datasets and then fine-tuned for specific domains, such as shipping, has shown promise in deciphering causal relationships. Although these models do not possess innate causal reasoning abilities, experimental results demonstrate that large language models, when fine-tuned, outperform existing causal algorithms in benchmark tasks related to causal discovery.

The true potential for AI to revolutionize shipping is contingent on creating solutions that are explainable, interpretable, and where there are strict protocols guiding how data is used to train and fine-tune models.

Mayflower Autonomous Ship and Lessons Learned

With the completion of the Mayflower Autonomous Ship voyage from England to United States, ProMare has proven that AI can completely and safely navigate a ship across the oceans autonomously – without crew or remote control (Kıcıman et al., 2023; MAS, n.d.).

This and other similar experiments will likely usher in a new era of fully autonomous vehicles at sea. You can quickly imagine, for example, a fleet of self-driving vessels trawling the oceans for weeks and months at a time, collecting data for marine science – from assessing the impact of climate change, improving weather predictions, monitoring sea life and marine chemistry, to supporting exploration of our deep-water ocean frontiers – all at a fraction of the cost and risk of outfitting a vessel full of sailors and scientists manually collecting that same data. At some point in the future, no doubt cargo will be transported across the oceans without any crew — just load it, point, and set it on its way.

However, a potentially larger and more immediate use of this technology is not for vessels that drive themselves all over the world, but rather to assist sea Captains in their own navigational and ship operations responsibilities. One of the lessons from the Mayflower project is that sometimes you do need a human on-board – to fix a broken manifold coupler to the auxiliary engine, or to repair the electrical connection that has frayed and causing a short. And with human lives at risk, you need humans that have the authority and responsibility for protecting those lives – for making the critical decisions necessary to fulfil their mission safely.

The MAS experiment provides valuable insight to guide how technology and AI can be used to revolutionize shipping:

- Data quality matters. Ships can encounter a vast array of conditions during their voyage. The Promare team used an off-the-shelf computer-vision algorithm from IBM to do some of its image segmentation and object detection, but to get the AI to classify those objects it had to develop a bespoke dataset. The team created a training set of millions of images of things that float in the sea, from buoys to boats to crab pots, taken in all kinds of weather and lighting conditions to train its algorithm (Kahn, 2021). Replicating this effort is clearly unsustainable. Instead, intelligent data must be married with powerful foundation models that are fine-tuned to different use cases. For example, Meta recently released "Segement Anything" a foundation model trained with over 1 billion masks on 11M images (Kirillov, 2023).
- There are advantages to treating the ship as a series of interconnected components. Each component fulfils a specific role: one handles computer vision from MAS's cameras, another analyses radar data, another utilizes GPS for navigation, and yet another optimizes power consumption among the various components drawing power from the battery. On top of this stack lies a decision-making engine responsible for determining the boat's course and controlling the engine throttle. The key benefit here is that this decision-making captain can continue functioning even if one of the underlying systems it relies on for perception and location data fails. Moreover, the

decisions made by the AI captain are explainable. Engineers have the ability to retrospectively review the inputs, their relative importance, and the resulting actions taken by the ship. Unlike a single neural network, it is not an opaque black box.

- Guardrails are vital: Adherence to essential regulations is crucial, as ships must always abide by important rules such as the International Regulations for Preventing Collisions at Sea (COLREGs) and the International Convention for the Safety of Life at Sea. When it comes to intelligent ship systems, they must consider real-time data to enhance decision-making processes while ensuring compliance with maritime regulations.
- Humans + AI solutions: The Mayflower AI Captain held a distinctive position, assuming complete control of the ship. However, the true potential of AI solutions lies in their capacity to assist ship captains and other stakeholders in processing vast amounts of data generated onboard. These AI tools can then suggest optimal decisions across various domains, including route selection, navigation, engine performance, HVAC systems, crew scheduling, and maintenance requirements. While AI can acquire contextual understanding and situational awareness, along with the algorithms to evaluate and determine effective options, it must still establish the most effective means of influencing decisions and executing actions. Ultimately, AI must earn the trust of its users, necessitating interaction with users, providing explanations for its reasoning behind certain decisions, and collaborating with users iteratively to arrive at the best possible choices.

Modern ships are characterized by their advanced technology, featuring an extensive array of sensors, devices, and servers that contribute to their day-to-day operations. To enhance the efficiency, sustainability, and reliability of the shipping industry, the integration of AI throughout the ship's operational systems is crucial. This entails incorporating AI into various aspects such as data acquisition, mechanical performance, logistics and supply chain management, port interactions, route selection, navigation management, and accounting for external factors like weather disruptions and geopolitical events. To create robust, reliable, actionable, and explainable decisions, a diverse set of routines are required. It is vital to ensure that the system can withstand IT outages, data contamination, model errors, and potential malicious actions. Additionally, each component must be designed to optimize human-machine interactions and ensure that the system's recommendations are easily understandable and applicable for stakeholders involved.

Foundation models have the potential to transform enterprise operations, showcasing impressive performance across various domains like natural language understanding, image generation, virtual assistants, data analysis, and healthcare. For the shipping industry, the crucial aspect is harnessing the power of these pre-trained models, adapting them to specific applications, and seamlessly deploying them either on-premises or in the public cloud. While foundation models introduce a new wave of AI capabilities, the underlying frameworks to enable them remain consistent. Considerations such as compute hardware, security-focused systems, intelligent workload orchestration, cloud-native applications, and robust development protocols involving DevOps, DataOps, and AIOps are essential factors to be addressed.

CHALLENGES AND RESEARCH DIRECTIONS

Rathore et al. (2022) review some of the challenges and future opportunities for edge computing, including:

- Security: The distributed architecture, mobility support, and data processing in edge computing raise security concerns. Thorough testing, access control systems, trust management, privacy measures, and intrusion detection systems are crucial. Security protocols, artificial intelligence, and encryption techniques can enhance protection, but fatal defects need to be addressed.
- Cost: Configuring, deploying, and maintaining edge computing devices can be expensive. Backup servers for fault tolerance and machine learning-based systems may incur additional costs. Developing cost-effective devices with low communication and computation expenses is essential.
- Power: Providing cloud-like remote services in remote areas requires high-power processors, voltage, and three-phase electricity, posing challenges.
- Data: Optimizing data usage is vital as edge computing devices process subsets of data, potentially overlooking additional data or wasting raw data. Effective data segregation and sharing across devices are crucial for efficiency. Computation-aware networking can enhance distributed systems.
- Heterogeneity: Resolving the heterogeneity of data and the ecosystem within edge computing networks is a challenge. Collaboration between different vendor systems, load balancing, synchronization, resource sharing, data privacy, and interoperability contribute to the complexity.
- Trust: User trust in edge computing systems is tied to security and privacy. Developing consumer trust models that foster confidence in adopting edge computing devices is an ongoing challenge.

Shipping stands out as a prominent example of edge deployment due to its unique characteristics, including restricted latency, inherent geographic distribution, autonomy demands, substantial data generation capacity, and the necessity for resilience, fault tolerance, and regulatory compliance. Further the shipping industry has ambitious decarbonisation ambitions which impacts all aspects of decision making (IMO, n.d.).

CONCLUSION

There have been many evolutions in computing from transistor computing to microprocessors, to networking and the internet. The past two decades have seen the rise of cloud computing and the many instances of these from public cloud to private cloud to hybrid and multi-cloud solutions. Edge computing is considered an extension of the cloud continuum because it complements and expands the capabilities of cloud computing. While cloud computing excels at handling massive data storage, complex data analysis, and resource-intensive tasks, edge computing focuses on processing data locally, enabling quick responses, and reducing the need for data transmission to the cloud. These two computing paradigms work together to provide a holistic and scalable solution.

By bringing computational power closer to the data sources, edge allows for faster processing and real-time decision-making, enhancing the overall efficiency of the system. The importance of edge enabled capabilities for shipping is obvious. Realizing the potential of edge computing is contingent on two factors:

- Extending cloud-centric development frameworks such as DevOps, DataOps, and AIOps to the Edge so that developers can seamlessly develop, deploy and improve new solutions within a single user framework.
- Leveraging new AI technologies such as foundation models rapidly on the edge. Deploying these powerful capabilities allows the industry to unlock sophisticated natural language understanding, image generation, data analysis, and decision capabilities at a fraction of the effort.

REFERENCES

Arnold, M. (2020). Towards automating the AI operations lifecycle. *Conference on Machine Learning and Systems*.

Bommasani, R. (2021). On the opportunities and risks of foundation models. *ArXiv Prepr. ArXiv210807258*.

Dell. (2023). *New Dell Edge Gateway 5200*. Dell. https://www.dell.com/en-ie/shop/gateways-embedded-computing/dell-emc-edge-gateway-5200/spd/dell-edge-gateway-5200/emea_edgegateway5200

Google. (n.d.). What is a Hybrid Cloud? *Google Cloud*. https://cloud.google.com/learn/what-is-hybrid-cloud

Hines, M. (2023). *What is the Future of Edge Computing?* Built In. https://builtin.com/cloud-computing/future-edge-computing

Hojlo, J. (2021). Future of Industry Ecosystems: Shared Insights & Data. *IDC Blog*. https://blogs.idc.com/2021/01/06/future-of-industry-ecosystems-shared-data-and-insights/.

IBM. (2021). *IBM Edge Application Manager*. IBM. https://www.ibm.com/docs/en/eam/4.1.

IBM. (n.d.). *What is AIOps?* IBM. https://www.ibm.com/topics/aiops

IMO. (n.d.). *UN body adopts climate change strategy for shipping*. IMO. https://imopublicsite.azurewebsites.net/en/MediaCentre/PressBriefings/Pages/06GHGinitialstrategy.aspx.

Iyengar, A. (2023). *Security at the Edge*. IBM. https://www.ibm.com/cloud/blog/security-at-the-edge

Kahn, J. (2021). What an autonomous ship named Mayflower can teach us about building better A.I. *Fortune*. https://fortune.com/2021/11/30/a-i-lessons-from-mayflower-autonomous-ship-eye-on-ai/

Kahn, J. (2023). The inside story of ChatGPT: How OpenAI founder Sam Altman built the world's hottest technology with billions from Microsoft. *Fortune*. https://fortune.com/longform/chatgpt-openai-sam-altman-microsoft/

Kıcıman, E., Ness, R., Sharma, A., & Tan, C. (2023). Causal Reasoning and Large Language Models: Opening a New Frontier for Causality. *ArXiv Prepr. ArXiv230500050*.

Kirillov, A. (2023). *Segment Anything*. arXiv, /arXiv.2304.02643. doi:10.1109/ICCV51070.2023.00371

MAS. (n.d.). *Mayflower Autonomous Ship*. https://mas400.com/.

Microsoft Azure. (n.d.). *Azure Stack Edge*. https://azure.microsoft.com/en-us/products/azure-stack/edge

Open Horizon. (2023). *Home*. https://open-horizon.github.io/

Pearl, J. (2009). *Causality*. Cambridge University Press. doi:10.1017/CBO9780511803161

Pearl, J., & Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect* (1st ed.). Basic Books.

Raghavan, S., & Shim, C. (2021). A new AI model could help track and adapt to climate change. *IBM Research Blog*. https://research.ibm.com/blog/geospatial-models-nasa-ai#fn-1

Rathore, V. S., Kumawat, V., Umamaheswari, B., & Mitra, P. (2022). Edge Computing: State of Art with Current Challenges and Future Opportunities. V. S. Rathore, S. C. Sharma, J. M. R. S. Tavares, C. Moreira, and B. Surendiran, (eds.). Rising Threats in Expert Applications and Solutions. Lecture Notes in Networks and Systems. Singapore: Springer Nature. doi:10.1007/978-981-19-1122-4_15

Scale Computing. (n.d.). *Intel NUC Enterprise Edge Compute Edition Built with Scale Computing*. https://www.scalecomputing.com/landing-pages/intel-nuc-enterprise-edge-compute-edition

Schabell, E. D. (2022). 5 reference architecture designs for edge computing. *Enable Architect.* Redhat. https://www.redhat.com/architect/edge-portfolio-architecture

Yusuf, K. (2023). Introducing watsonx: The future of AI for business. *IBM Blog*. https://www.ibm.com/blog/introducing-watsonx-the-future-of-ai-for-business/