Palantir

Edge AI + Autonomy: Model Deployment, Operationalization, & Edge AI

December 2022
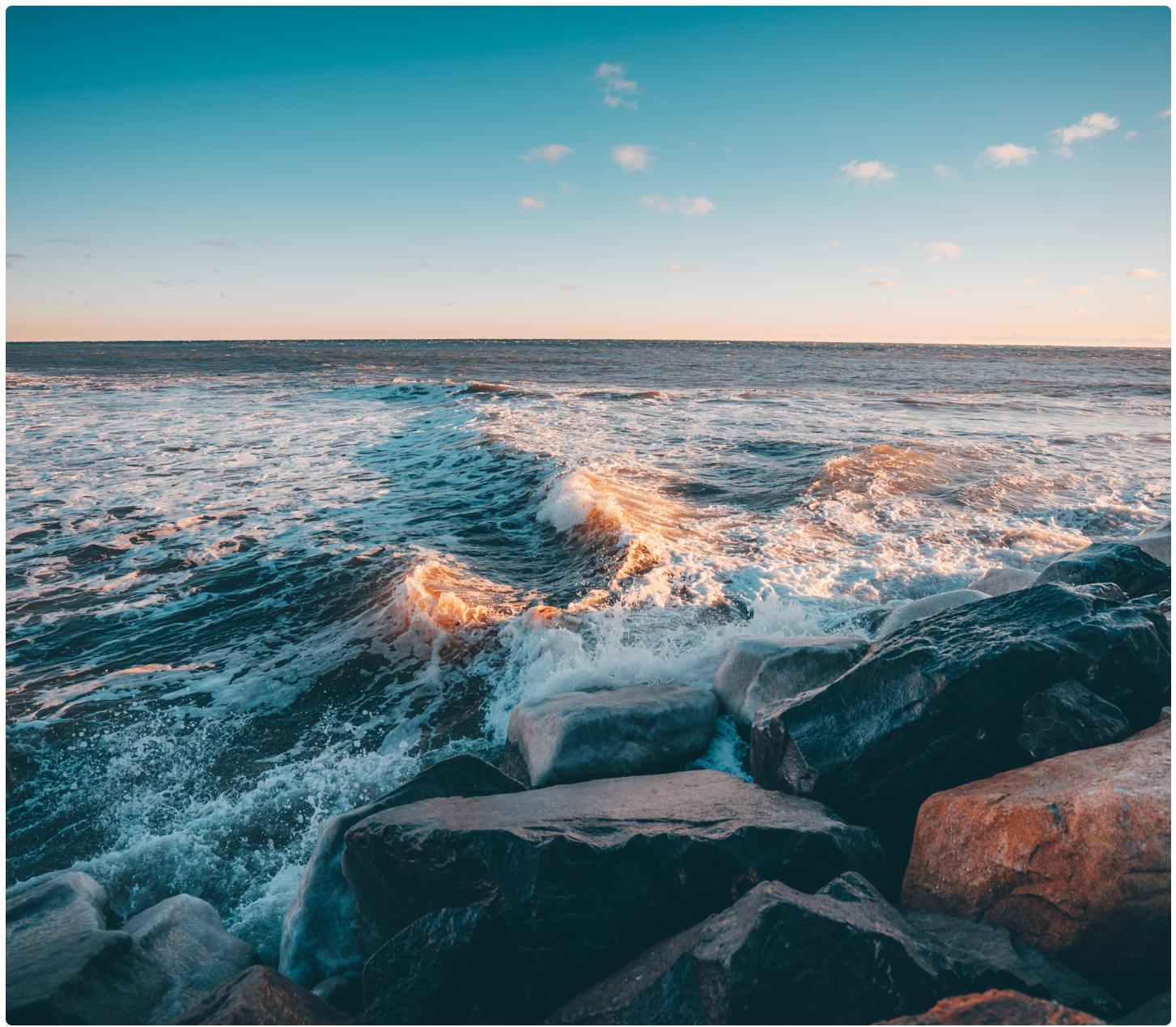
The Challenge ↘

Through our work supporting forward operations for customers across every branch of the Department of Defense and several stakeholders within the Intelligence Community, Palantir has discovered AI challenges common to nearly all edge deployments in contested, disconnected environments. These challenges are particularly acute within the maritime context and, without effective mitigation, will limit Navy efforts to develop effective autonomy capabilities.
Common challenges include:

- ### Bandwidth Constraints for Data Backhaul:

Operating at the edge, particularly in Delayed/Disconnected, Intermittently-Connected, Low-Bandwidth (D/DIL) environments, comes with significant bandwidth constraints. As a result, operators are frequently limited by the network when it comes to which sensors and platforms they are able to use, and how they are able to use them. While tactical communications pathways can overcome contested communications and transmit limited subsets of high priority data back for human review, analysis, and mission enablement, the underway systems often lack a solution for identifying which subsets of data are highest value and worthy of priority backhaul.

- ### Model Overgeneralization:

Too often, model developers take a "one size fits all" approach to model development because they assume a single platform will only be able to operate a single model and seek to train that model to all environments in which that platform may operate. This results in a "jack of all trades, master of none" performance characteristic profile and doesn't fully leverage the power of finely tuned, highly differentiated AI modeling.

- ### Algorithm Deployment Limitations:

Navy platforms span a wide range of system architectures and software versions, making it difficult to deploy new models across the fleet in a streamlined way. All too often, developers must manually upload models and updates to individual platforms while lacking comprehensive visibility into which models are deployed to which platforms at any given time. This results in frequent setbacks to the build, test, iterate, and retest process and makes it difficult to meet evolving priorities at mission speed.

The Solution ↘

Palantir can provide capability to streamline model deployment and operationalization in support of edge AI. With the ability to conform to nearly any size, weight, and power (SWAP) requirements, Palantir's software can be loaded onboard a variety of heterogeneous platforms to form a comprehensive solution linking the centralized ML operations and model refinement environment with tactical operations at the edge. Within this data ecosystem, Palantir can enable the Navy to best leverage high scale sensor and operational data to drive superior outcomes — both in short term tactical operations, and long-term model development strategy. Palantir's key benefits for model deployment and operationalization include:

- ### Priority Data Backhaul:

Palantir's solutions can help the Navy solve its tactical bandwidth problem by managing existing data slated for transfer in D/DIL environments. Our software can serve as the intermediary between onboard models and the onboard tactical communications pathways responsible for transmitting priority information. Data tagged as high priority through onboard processing can be queued and sent in near-real time to several receiving systems, including nearby vessels, systems, and tactical nodes. In a sample workflow, an unmanned platform equipped with EO/IR sensors and running an onboard object detection algorithm could send a still image of a potential detection cut from the feed alongside coordinates to a human operator stationed at a nearby node for confirmation. All operational data not sent through these pathways can be backhauled once the mission is complete and bandwidth is plentiful. By pushing only quality transmissions, our software frees tactical bandwidth while providing greater strategic value.

- ### Right Model, Right Platform, Right Time:

Palantir software is installable on any number of Navy-selected form factors, including third-party devices, existing vessel systems, and Palantir-provided hardware. Once onboard, Palantir's edge AI capability can dynamically cycle between competing models to leverage the best performing model for the specific operating environment. For example, rather than trying to write a model that performs equally well in choppy and smooth seas, Navy modelers can leverage our software to enable a platform to toggle between models tailored to deliver the best performance in whichever particular sea state the platform is operating. These model swaps occur in real time, with zero downtime, to ensure platforms are maximally equipped to meet their mission.

The Solution ↘

Rapid Model Delivery to the Edge:

Palantir's software brings continuous integration / continuous delivery (CI/CD) best practices to model deployment, enabling streamlined delivery from a centralized model testing and iteration environment to the tactical edge. Our software connects the model catalog, test and evaluation, and the decision and release cycle with model deployments to heterogenous platforms. Our solutions can remotely deploy model updates to various platforms worldwide alongside a single dashboard to centrally manage and monitor deployments, offering Navy developers advanced visibility into which versions of which models have been positioned on which platforms, as well as tooling to push or recall new models and versions to any vessel across the fleet.

All Palantir software is built to embrace openness, modularity, and interoperability. We design solutions to enhance the heterogeneous systems and capabilities already at the Navy's disposal, rather than replace these capabilities outright. In line with this mission, our solutions leverage open APIs and open data formats to enable efficient operations with a wide array of models and tools developed by both Navy and third-party stakeholders.