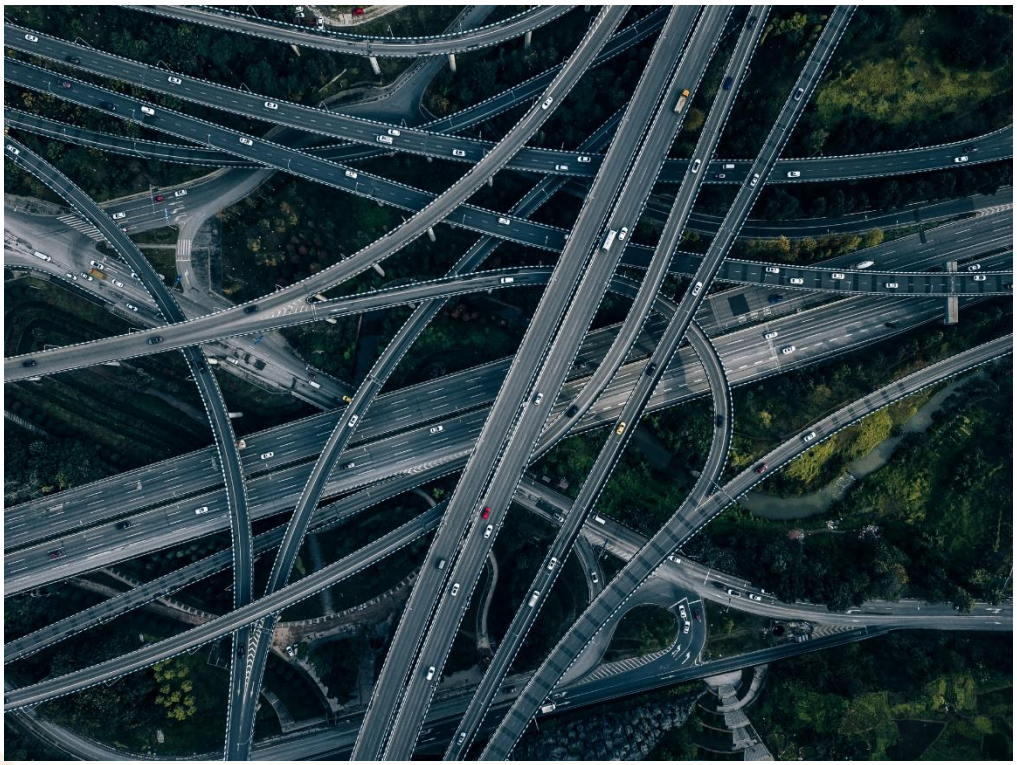
In collaboration with:

NAVIGATING THE EDGE: HOW TO SUCCESSFULLY BUILD DISTRIBUTED SOLUTIONS

Edge computing has become a popular choice among solution developers looking to meet stringent application requirements like ultra-low latency. This report explores the deployment strategies of six companies and extracts lessons for others on how to build a successful solution leveraging edge computing.

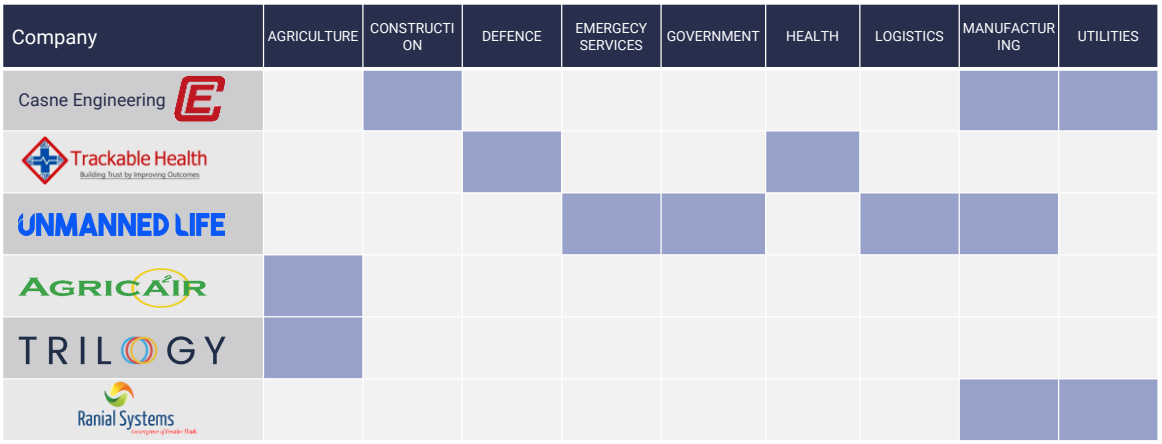


Kuba Smolorz, Consultant | kuba.smolorz@stlpartners.com | April 2023

Executive Summary

Edge computing is creating new business opportunities for solution providers who can take advantage of its performance characteristics. However, deploying applications across a distributed and heterogenous environment, as opposed to the single, consistent environment of centralised cloud creates complexity. There is much to be learnt from the pioneering solution providers who have already had experience navigating these challenges. In this report we bring together the findings from six of these solution providers (see Figure 1).

Figure 1: This paper focuses on lessons from six early adopters of edge computing



While their solutions, end customers and approach to edge deployments vary, there were some universal challenges they all had to overcome. These serve as important lessons for others looking to deploy at the edge and are summarised in Figure 2.

Figure 2: Key recommendations to mitigate challenges of deploying at the edge

Key challenge of deploying at the edge	How to overcome this challenge
Edge applications often require real-time processing of data from many different sources	**Ensure analytics software can handle data heterogeneity:** Solution providers should focus on being able to standardise and filter data at the edge to handle disparate incoming data from a variety of different devices, in different formats/structures, shared with different frequencies.
Deploying applications and updates across a distributed environment	**Ensure the development platform used is designed for distributed applications:** Platforms will be able to reduce the complexity of setting up and managing applications with workloads running in many different locations. **Adopt a modular and portable application architecture:** Edge applications should be designed as a set of independent modules or microservices that are flexible and can be independently modified, updated or scaled.
Diversity of edge devices	**Ensure bespoke integration is minimised where possible:** Solution providers should ensure that repeatable blueprints are created wherever bespoke device integration is completed which can be used to assemble future applications.
Edge devices are often deployed in remote or harsh environments	**Develop for intermittent connectivity:** Edge applications should be designed to handle intermittent connectivity and operate offline by ensuring critical application workloads still operate even if connectivity drops.

© STL Partners

Table of Contents

© STL Partners

Table of Figures

Introduction

Solutions providers used to be faced with an imperfect dilemma: either deploy your application in the hyperscale cloud or install your application directly onto the device. Over the last few years, however, edge computing has provided an attractive alternative – it combines the scalability and flexibility of cloud with ultra-low latency, data sovereignty, and reduced data transfer costs.[1] These attributes can enable new, and improve existing, applications such as drone inspection of assets or massive industrial IoT processing.

However, solutions providers looking to deploy at the edge face challenges, particularly in moving beyond trials and POCs to scaled deployments. These challenges can range from customer and market challenges (e.g. a lack of maturity of understanding from customers on the business benefits of deploying at the edge) to technical challenges (e.g. how to architect the application to run effectively in a distributed environment). We have identified four key challenges that solution providers encounter while deploying at the edge:

1. **Deploying applications and updates across a distributed environment** as opposed to a single location e.g. a centralised cloud. In a centralised cloud deployment, resources are typically unlimited, stable, and consistent, which makes it easier to optimise application performance. However, in a distributed environment, there is a high degree of variability making it challenging to automate the testing, deployment and management of applications across all edge devices while ensuring the applications are performant.

2. **Edge applications often require real-time data processing** to deliver insights and decision-making capabilities but equally, to support mission critical business automation. Processing data in real-time is challenging and requires advanced computing and analytics capabilities.

3. **Diversity of edge devices,** which have different operating systems, hardware, and software configurations. Edge applications must be designed to be compatible with a wide range of devices and must be able to adapt to different environments and network conditions.

4. **Edge devices are often deployed in remote or harsh environments**, such as factories, oil rigs, and farms, which often have limited or unreliable network connectivity and will not be able to provide datacentre like conditions.

In this report, we will focus on technical challenges and explore how to limit the complexities of deploying and managing edge applications and open path for innovation. Through our conversations with six edge solution providers targeting a range of industry verticals, we delve into the intricacies of deploying and managing edge applications[2], identify the obstacles that solution providers face, and explore ways to overcome them.

[1] To learn more about the basics of edge computing, see 'What is edge computing?'
[2] With a particular focus on device or on-premise edge computing

Edge for IIoT applications

Casne Engineering, company overview

Casne Engineering is an engineering, systems integration and technology services firm which works on bespoke engagements with clients across industries like manufacturing, construction, and utilities.

How Casne Engineering benefits from edge computing

Casne uses on-premise edge computing servers to provide localised processing and storage for industrial IoT (IIoT devices). These devices can generate vast amounts of data, though much of this will not be useful as it will reflect the normal functioning of an asset. Edge computing enables Casne to aggregate and filter data close to its source reducing the amount of data that needs to be transmitted to a centralised cloud, saving cost (see Figure 3). This is particularly valuable for scaled solutions with large numbers of sensors and therefore greater volumes of data.

> **"The main benefit of edge is being able to process data close to its source, rather than sending it all to the cloud for processing".**
>
> *Technical Program Manager*

Figure 3: Key benefits of edge computing for Casne Engineering

Source: STL Partners

One of Casne's projects which exemplifies the benefits of edge computing uses IIoT sensors to monitor building systems like heating, ventilation and air conditioning (HVAC) and lighting. The sensors validate that the automation has been correctly configured during the construction process and identifies any defaults. Traditionally, an on-site engineer would physically check the sensors, which is time consuming and costly as well as restricted by the number of engineers employed. Casne offers a cloud-based dashboard which collates inputs from the edge devices. Using this, the engineers have a complete view of asset performance and as sensor data is processed as it's produced rather than storing it to process in batches, engineers are alerted to important developments in real-time.

© STL Partners

Lessons learned: How Casne overcame key challenges associated with edge deployment

The most important challenge Casne has overcome is standardisation and it has manifested itself in two ways:

1. **Lack of standardisation in the IoT industry** means proprietary devices operate in different ways and Casne's analytics software had to be designed bespoke for each type of sensor. Building its software on an edge platform has allowed Casne to move to a more modular application architecture where independent segments of software can be reused for different sensor types. While some bespoke design is still needed, the entire application doesn't have to be built from scratch reducing development time. Additionally, proprietary devices often use different communication protocols and data formats meaning data must be standardised prior to analysis. Casne works with a data platform partner to achieve this.

> **"When you have data coming from lots of different sources, it all needs to be standardised before it can be analysed"**
>
> *Technical Programme Manager*

2. **Deploying and updating edge applications**, which can be difficult with proprietary hardware due to compatibility issues. Casne navigates this challenge by deploying its software in a containerised environment. Containers provide a standardised and isolated environment that can be updated and tested independently of the underlying host system. This makes it easier to manage updates across different edge devices and reduces the risk of compatibility issues or conflicts with other software.

> **"Deploying and updating applications at the edge used to be a nightmare, but using a containerised environment has reduced this complexity"**
>
> *Technical Programme Manager*

© STL Partners

Edge for wearable health devices

Trackable Health, company overview

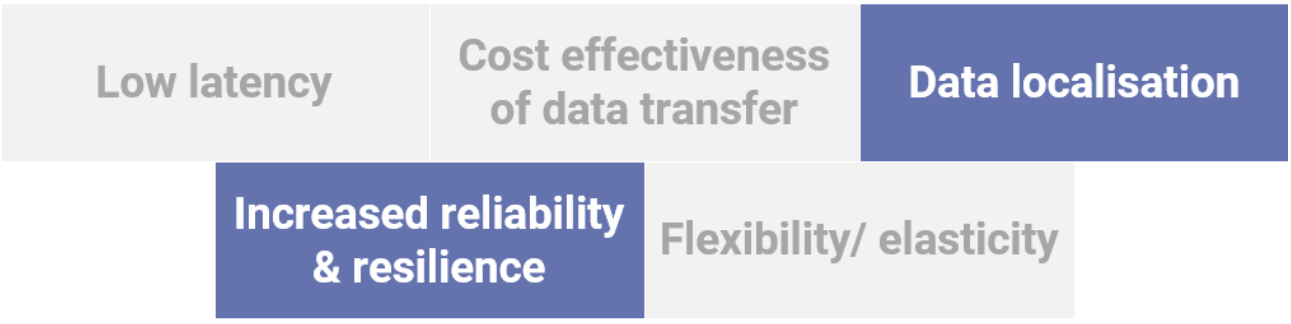
Trackable Health provides a hardware agnostic platform for wearable devices which track the fitness of its users.

How Trackable Health benefits from edge computing

Trackable Health works extensively with military personnel, enabling soldiers to customise their training and track the environmental conditions that might impact them. This means that the upmost concern for Trackable Health is maintaining data security (see Figure 4) which is bolstered by analysing all data on-premises using edge servers. If certain data from the wearables e.g. location, was accessed by an unauthorised party, the safety of military personnel could be jeopardised and therefore Trackable Health cannot risk sending the data to the cloud. Resiliency is another key benefit of edge computing as military training/operations are often conducted in remote areas with no or limited connectivity. This means data needs to be processed and stored on-premises rather than relying on remote compute resources.

> **"Security is without a doubt the biggest benefit of edge as it allows sensitive data to remain on the customer premises rather than being sent to a central cloud"**
>
> *Founder*

Figure 4: Key benefits of edge computing for Trackable Health



Source: STL Partners

Lessons learned: How Trackable Health overcame key challenges associated with edge deployment

The main challenges of scaling edge deployments for Trackable Health comes from the complexity of updating and deploying applications in distributed compute environments.

1. **Deploying application updates in distributed environments** is challenging, for example, ensuring the right updates are applied to the right edge endpoints or deploying updates without having to take the application offline. Trackable Health leverage an edge platform which reduces some of the complexity and enables capabilities like dynamic application updates with minimal downtime.

> **"Deploying applications in distributed environments is very challenging as applications will run on a variety of different devices, infrastructure and networks"**
>
> *VP, Sales Engineering and Services at a supporting edge platform*

2. **Trackable Health have also adopted a loosely coupled microservice architecture** in which the microservices communicate via an intermediary message broker and are therefore independent of each other. This type of architecture makes the application easier to maintain, upgrade and integrate.

3. **Ensuring secure communication between the edge client and applications** is critical for maintaining the confidentiality, integrity and availability of data and applications. This is especially challenging due to the large cyber-attack surface area caused by many edge devices. When an edge device wants to access an edge server, that request is first intercepted by an intermediary which can evaluate the request to prevent any unauthorised access or malicious activity. To improve the security of edge deployments and data utilisation, Trackable Health is also adhering to the global healthcare industry standards for health record data interoperability.

© STL Partners

UNMANNED LIFE

Edge for autonomous robotics

Unmanned Life, Company Overview

Unmanned Life is an autonomous robotics solution provider for industrial applications. The solutions involve orchestrating lots of singular robots into fleets to deliver certain use cases e.g. drones to monitor large swathes of forest or AGV's to move materials around a warehouse.

How Unmanned Life benefits from edge computing

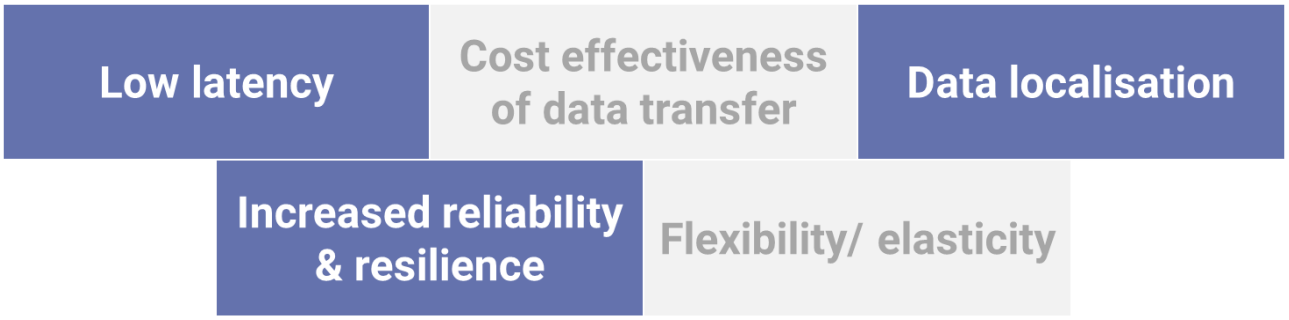
Unmanned Life's delivery of security critical applications require the low latency, resilience and data sovereignty benefits of edge; however, edge compute resource is limited and relatively expensive. Application workloads are therefore split up and ran in different locations depending on their technical requirements, only necessary workloads run at the edge (Figure 5 summarises the benefits of edge for Unmanned Life). The distribution of workloads can be grouped into two models:

> **"Application workloads are split up and run in different locations depending on their technical requirements e.g. workloads that require low latency run on the customers premise "**
>
> *CEO*

1. **Network edge:** Some application workloads run on servers at the customers' premises while others run in centralised clouds or at the network edge. For example the workloads which control a fleet of drones are latency sensitive and run on-premise while processing the images captured by the drones isn't latency sensitive and can run in datacentre.

2. **On-premise edge:** In areas with poor/ no connectivity, entire applications will run on-premises. This includes a private network creating a completely independent system.

Figure 5: Key benefits of edge computing for Unmanned Life



Source: STL Partners

Lessons learned: How Unmanned Life overcame key challenges associated with edge deployment

The main challenges that Unmanned Life has overcome when deploying at the edge are:

1. **Running applications across different environments** to meet customer and application requirements make it crucial for applications to be highly portable and perform equally well regardless of the underlying infrastructure. Unmanned Life deploys its applications in containers meaning applications don't have to be re-engineered for each environment type.

> **"Our applications are built using microservices deployed in containers, so they are highly portable and performant across any underlying environment"**
>
> *Cloud & Networks Architect*

> **"We need to test the application in similar conditions to the production environment. We can use simulations but for each major release, we need to test on the actual environment"**
>
> *Cloud & Networks Architect*

2. **Regular updates across a distributed environment.** Unmanned Life have commitments and, in some instances, legal obligations to customers and partners to regularly update applications. DevOps methodologies such as CI/CD pipelines can help to achieve this at scale while doing quality assurance checks before any update goes live. Implementing CI/CD for edge applications is more challenging than cloud applications due to the naturally heterogeneous nature of edge. For example, applications need to be tested in an environment that closely resembles the actual deployment environment. To combat this, Unmanned Life uses simulations to mimic the final environment as closely as possible.

3. **Dynamic distribution of workloads across the cloud and edge** to optimise the use of limited edge resource. Ultimately, Unmanned Life is aiming to dynamically distribute workloads to meet their requirements with the optimal resource and infrastructure utilisation rather than predetermining where workloads should run before deployment. Close collaboration with hyperscaler partners and telecoms operators is key to unlocking this capability.

© STL Partners

Edge for video analytics

Agricair, Company Overview

Agricair provides real time video analytics solutions for the dairy industry for monitoring purposes including animal welfare, as well as to optimise general farm operations e.g. feeding times.

How Agricair benefits from edge computing

A typical Agricair deployment involves 20-30 high quality cameras (4k) that provide continuous video streams which are analysed in real-time, flagging any notable events to the customer e.g. animal fallen over. Due to poor connectivity on farms its unfeasible to upload all video to the cloud for processing which is why it's analysed using on-premise edge servers to sift through the footage, most of it being routine and identify, flag and upload video clips of notable events to the cloud. Important video clips will be stored on the edge server in case of lost connectivity, until they can be uploaded to the cloud. This increased resiliency is crucial as even a second of lost video footage can be vital, for example, a clip of an animal being mistreated. Finally, data localisation is a big consideration. Some of these clips may capture sensitive events. Customers therefore want to have a solution where the video is kept and processed on-premises, the relevant clip is packaged and encrypted before being securely uploaded to the cloud. Figure 6 summarises the benefits of edge for Agricair.

"Poor connectivity in rural areas means it's not feasible to upload all our video to the cloud, processing data on-premise means we can upload only the most crucial data to the cloud."

Technical Director

Figure 6: Key benefits of edge computing for Agricair

Source: STL Partners

Lessons learned: How Agricair overcame key challenges associated with edge deployment

An overall barrier to Agricair's solution is related to the limited bandwidth and availability of connectivity in rural areas in the US which will only become a greater issue as camera quality increases,

producing greater volumes of data. On-premise edge computing will be essential to handling this and so it has been essential that Agricair overcame the following challenges with edge deployments:

1. **Software development, testing and roll out across a distributed environment.** Cloud applications can be developed and tested for quality assurance and then deployed to a single location i.e. a cloud server. Edge brings in a whole new set of complexities due to its heterogenous and distributed nature meaning ensuring an application will be performant across all customer deployments is much more difficult. Agricair deploys its application in a containerised environment to make sure the environment within each container is consistent regardless of the underlying infrastructure so the application works the same in testing and customer implementation.

2. **Application development can be time consuming** but by building their application on an edge platform, Agricair has been able to benefit from low-code programming, allowing them to reduce the time. A simplified development process has enabled a greater focus on their core product and improved their offering.

> **"Using an edge platform has enabled us to reduce time spent on application development like building the front-end, which means we are able to focus more on our solution."**
>
> *Technical Director*

> **"We will continue to be limited by the lack of reliable connectivity with high bandwidth in rural areas."**
>
> *COO*

3. **Operating in low connectivity environments** such as farms means edge devices may not always be connected to the internet, which can make it difficult to deploy updates in real-time. To combat this, updates are uploaded to and stored on the on-premise server and the device queries the edge server every couple of hours to see if there is an update. If there is, it will pull the update automatically.

Edge for smart agriculture

Trilogy Networks, company overview

Trilogy Networks provide end-to-end solutions to the agriculture industry which includes network and IT infrastructure provisioning, application integration and device lifecycle management.

How Trilogy Networks benefits from edge computing

Trilogy Networks separates out the workloads which make up an application and deploys them in different locations, depending on their technical requirements. For instance, some workloads require low latency e.g. a fruit picking robot analysing video streams to decide if fruit is ripe enough to pick. This decision needs to happen in milliseconds, so the video needs to be processed using on-premise edge servers. Other applications do not have such requirements, such as the analysing images from a drone flying over long distances, meaning this data can be processed in a network edge data centre.

"Every use case has different technical requirements which dictate where the workloads run e.g. on-premise or in a regional edge data centre"

Co-Founder

Trilogy Networks leverages edge for resilience which is particularly important as updating applications in areas of poor connectivity e.g. rural areas can be challenging. With edge computing, Trilogy Networks is able to store the update on-premise until sufficient connectivity is available for the update to be downloaded and installed. Figure 7 summarises the main benefits of edge computing for Trilogy Networks' different applications.

Figure 7: Key benefits of edge computing for Trilogy Networks

Source: STL Partners

Lessons learned: How Trilogy Networks overcame key challenges associated with edge deployment

The main challenges that Trilogy Networks has had to overcome when deploying at the edge are:

> **""Working with partners that can provide data harmonisation is key as we're able to simultaneously analyse data from lots of different sources"**
>
> *Co-Founder*

1. **Extracting business insights from many disparate data sources** calls for data harmonisation before analysis can begin. Trilogy Network's solution often uses many different sensors and data points which need to be standardised. Trilogy Networks rely on a data harmonisation partner, with a specialised focus on agriculture, that leverages its vertical expertise to synthesise unstructured, and often abundant, data into key actional business insights. Furthermore, the partner's agriculture-specific datasets e.g. satellite imagine, can augment data collected by Trilogy Networks' sensors for a more comprehensive and effective approach.

2. **Maximising the usage of limited edge resource** can be helped by activating compute on demand. On-premise edge servers can be kept offline or designated for other functions until they are needed. To continue with the fruit picking robot example, the compute that controls the picking function of the robot can be inactive until the video analytics function determines that the fruit is ripe enough to be picked. Once the ripeness check has passed, the compute that controls the picking function is activated, the action is carried out and the function is deactivated. This negates the need to over provision compute.

© STL Partners

Edge for smart grid

Ranial Systems, company overview

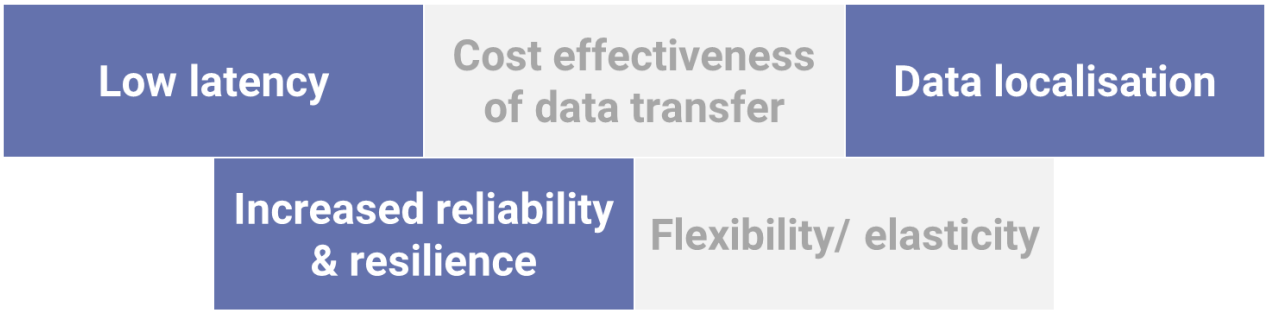
Ranial Systems is an IoT solution provider for the energy and manufacturing industries. In this case study we focus on Ranial Systems' support for the energy sector, particularly in enabling autonomous operations of distributed energy resource systems e.g. solar panels.

How Ranial Systems benefits from edge computing

Ranial Systems uses on-prem edge servers to process data in real-time and enable automation of assets in the energy sector e.g. controlling a battery to autonomously discharge if there a dip in solar energy production. An increased focus on renewables makes for a more distributed energy production system which adds complexity, necessitating more automation to operate a greater number of distributed assets. Automated control of power grid assets is mission critical, downtime or errors are not tolerated, meaning it requires ultra-reliability, low latency and data sovereignty, all key benefits of edge computing over processing in the cloud (see Figure 8).

> **"You cannot run mission critical applications in the cloud due to reliability and latency issues"**
>
> *CEO*

Figure 8: Key benefits of edge computing for Ranial Systems



Source: STL Partners

Lessons learned: How Ranial Systems overcame key challenges associated with edge deployment

The challenges Ranial Systems faced when deploying at the edge are:

> **"Achieving interoperability in modular designs has always been a challenge for edge computing"**
>
> *CEO*

- **Achieving interoperability between modular components**. As edge computing involves a vast array of devices, sensors, and platforms, there is a high likelihood that these components were developed using different technologies, protocols, and standards. This makes it challenging to connect these modular components and ensure they can communicate with each other effectively which is particularly important for Ranial Systems to ensure the smooth and efficient operation of mission critical systems like control of power grid assets.

- **Analysing data in real-time** i.e. as soon as it's generated requires a streaming architecture. Crucially this allows for automation as changes can be made almost instantly in response to any developments e.g. a battery is discharged in response to a dip in power production from a generator. This also allows key stakeholders e.g. engineers, to be informed of the most important developments as key insights are uploaded to the cloud and a notification is sent out to the relevant stakeholder as soon as they occur. Additionally, Ranial Systems consistently enhances its AI models through continuous retraining by incorporating the latest data.

- **Ensuring an ultra-reliable system** requires Ranial Systems to use a mesh network topology where each edge device in the network is interconnected using connectivity like LoRa or Zigbee. This network topology ensures a highly resilient and robust network since each device can communicate with several other devices, creating multiple paths for data to travel. This redundancy means that if one path is blocked or congested, data can still reach its destination through another path.

© STL Partners

Conclusion and recommendations

In order to reap the benefits of edge computing, key challenges need to be overcome that arise from working in a distributed, complex and heterogenous environment. From our conversations with solution developers, we have identified four key challenges that they face and below we summarise effective strategies to overcome them (see Figure 9).

Figure 9: How solution developers can overcome edge deployment challenges

Key challenge of deploying at the edge	How to overcome this challenge
Edge applications often require real-time processing of data from many different sources	**Ensure analytics software can handle data heterogeneity:** Solution providers should focus on being able to standardise and filter data at the edge to handle disparate incoming data from a variety of different devices, in different formats/structures, shared with different frequencies. Solution providers will likely need to combine vertical expertise (e.g. to prioritise data protocols most commonly used by their end customers) with data management and analytics expertise to build a strategy that can handle demanding, real-time scenarios.
Deploying applications and updates across a distributed environment	**Ensure the development platform used is designed for distributed applications:** Platforms designed to support edge computing deployments will be able to reduce the complexity of setting up and managing applications with workloads running in many different locations. Platforms that have pre-built tooling and dashboards that provide real-time visibility on application performance can help solution providers to scale more easily. **Adopt a modular and portable application architecture:** Edge applications should be designed as a set of independent modules or microservices which can be run in different locations depending on their technical requirements e.g. microservices responsible for latency sensitive workloads can run on a customers premises. These modules should be easily portable meaning they will be performant, regardless of the underlying environment they are running on. Finally, modules can be designed to communicate through Pub/Sub messaging (also known as event driven architecture) to support scalability, reliability, and enable loose coupling. This further adds to the modularity of the application as modules operate independently of each other and therefore can be scaled, modified or updated without affecting the rest of the system.
Diversity of edge devices	**Ensure bespoke integration is minimised where possible:** Solution providers should ensure that repeatable blueprints are created wherever bespoke device integration is completed which can be reused to assemble future applications. This ensures that solution providers can productise their offerings as much as possible, particularly for common device types.
Edge devices are often deployed in remote or harsh environments	**Develop for intermittent connectivity:** Edge applications should be designed to handle intermittent connectivity and operate offline or with limited network resources. Critical application workloads should run on-premise or on-device to guarantee business continuity even if connectivity drops. Workloads that provide non-critical, value add features can be reliant on connectivity, with the ability to bulk upload data if connectivity is lost and re-established.

© STL Partners

A message from our sponsor

Vantiq is helping innovators across industries, including agriculture, real estate, telcos, defence, and healthcare, to create real-time edge-to-cloud solutions without the risks of time, cost, or complexity. Built from the ground up with distributed computing in mind, Vantiq provides the leading low-code platform to enable rapid development and deployment of large-scale distributed applications that orchestrate how IoT devices, camera systems, AI algorithms, and humans can drive smarter business outcomes. From disaster response systems, to connected patient monitoring, to the world's smartest building, the Vantiq platform is enabling companies across the globe to realize the benefits edge computing has to offer and how to navigate their journey to the edge.

For more information, visit **https://vantiq.com/** or contact **info@vantiq.com**

© STL Partners

STL PARTNERS

Research

Consulting

Events

www.stlpartners.com