T-DAB.AI

SMARTER DATA, BETTER DECISIONS

Gold
Microsoft Partner
Microsoft

OUR MISSION

T-DAB.AI's mission is to radically accelerate the building of Industrial IoT (IIoT) powered businesses through decentralised AI

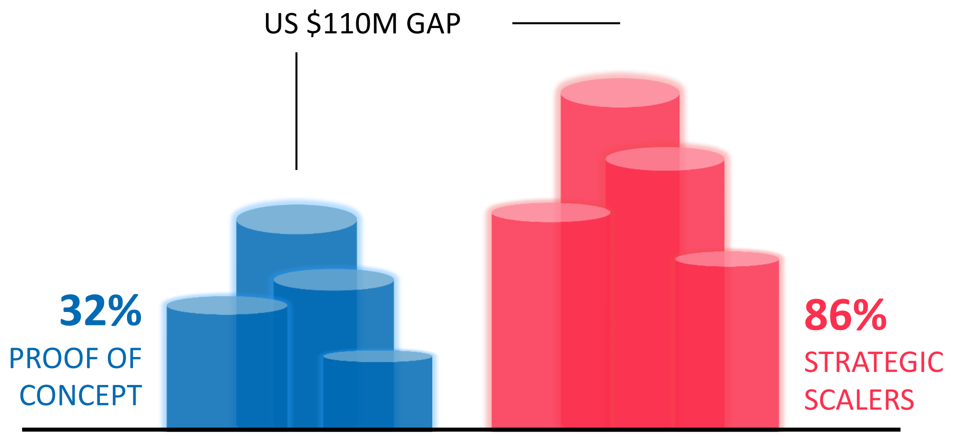
AI CUSTOMER OPPORTUNITY

Predicting and Diagnosing Machine Health using AI: **Manufacturers that fail to do this at scale are on average $110m worse off**

Predicting failures to maximise uptime and yield, while minimising wastage and carbon footprint is a ubiquitous AI application for industry

An estimated value for manufacturers is $500-700bn

84% of companies failing to scale AI



The difference in return on investments between companies in the Proof of Concept stage and Strategic Scalers equates to an average of US $110 million.

50% of IT leaders struggle to move their ML/AI projects past PoC

Acute problem for the 'Mighty Middle' (£10m-£100m revenue)

The opportunity is to close this ROI gap

CASE STUDY

Case Study: Predicting Future Failures and Detecting Anomalies With Edge Auto-ML

Predictive Maintenance And Quality Assessment: Predicting Machine Jams And Low Quality Outputs

PROBLEM

Spoilage results in wasted material, increased downtime, and inefficient power use increasing carbon footprint

SOLUTION

A suite of ready to consume ML pipelines built on top of a scalable IIoT platform in order to predict the risk of and time to spoilage events

5% REDUCTION OF SPOILAGE EVENTS CAN INCREASE REVENUE PER LINE BY +£3M

USE CASE OPTIMISED ML PIPELINES >90% ACCURACY

THE GOAL

To accurately predict where and when machine failures will occur so that…

We can take an action to minimise the chance that the machine will fail, thus preventing stoppages and lost productivity through down time

Predict failures and stoppages to prevent them = Maximise up time

Diagnose and locate abnormal function = Minimise downtime through more targeted, efficient, and well planned maintenance and ordering of parts

Intelligently optimise control = Dynamically optimise the trade-off between failure/stoppage risk and productivity

THE MANUFACTURING PROBLEM

Identification of Tear-off: Affecting Actual operating time

Produce a number of cans by employing a number of bodymakers

Each bodymaker has the tooling installed to produce can bodies to from a coil of aluminium

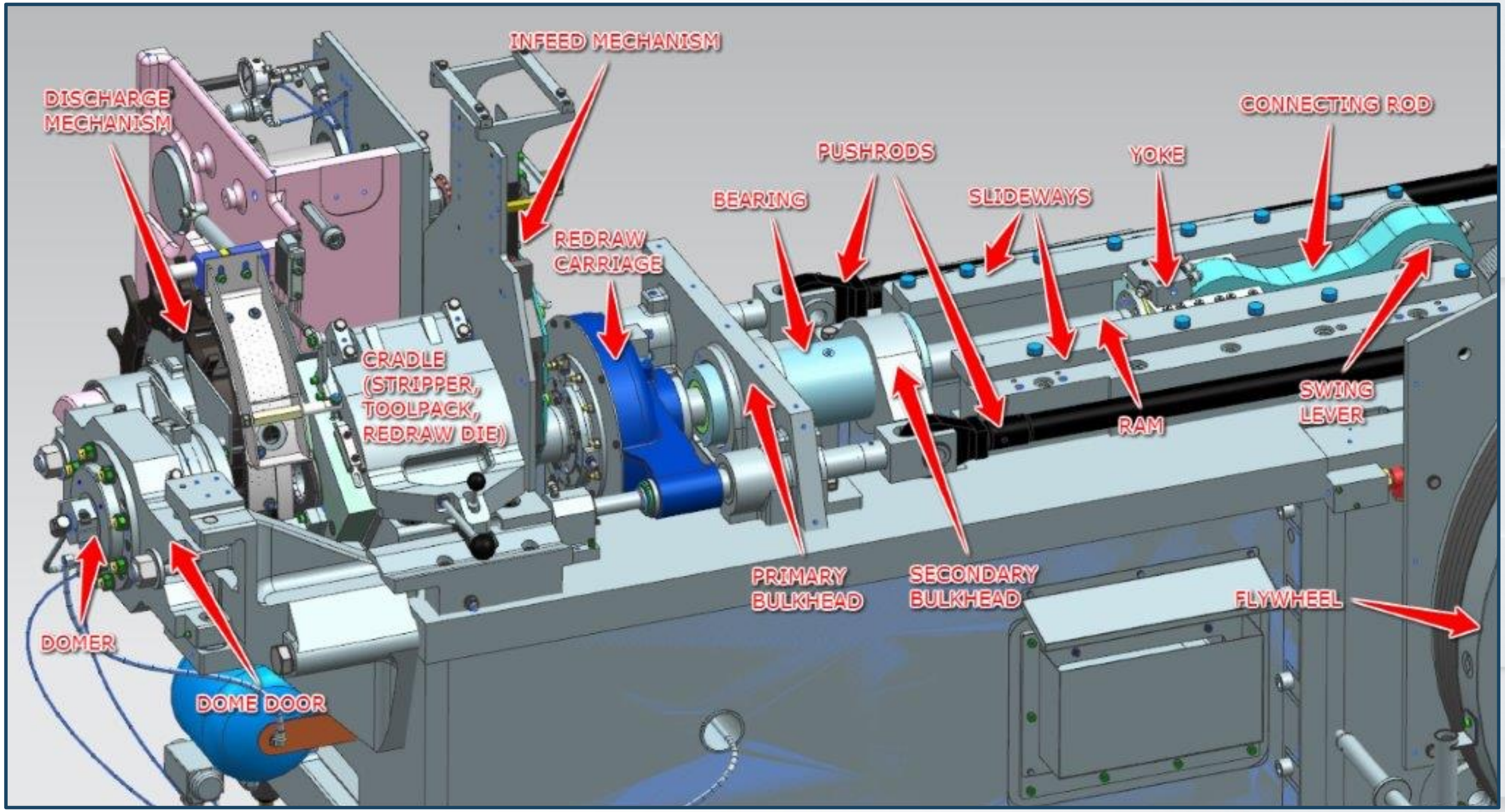
Each bodymaker loads a single cup at the beginning of each stroke, this cup will be ironed into a can body, which will be stripped off the punch and discharged via a conveyor. The punch will retract ready for the process to begin again.

A human operator sets the speed manually, at higher speeds the chance of a Tear-off is greater

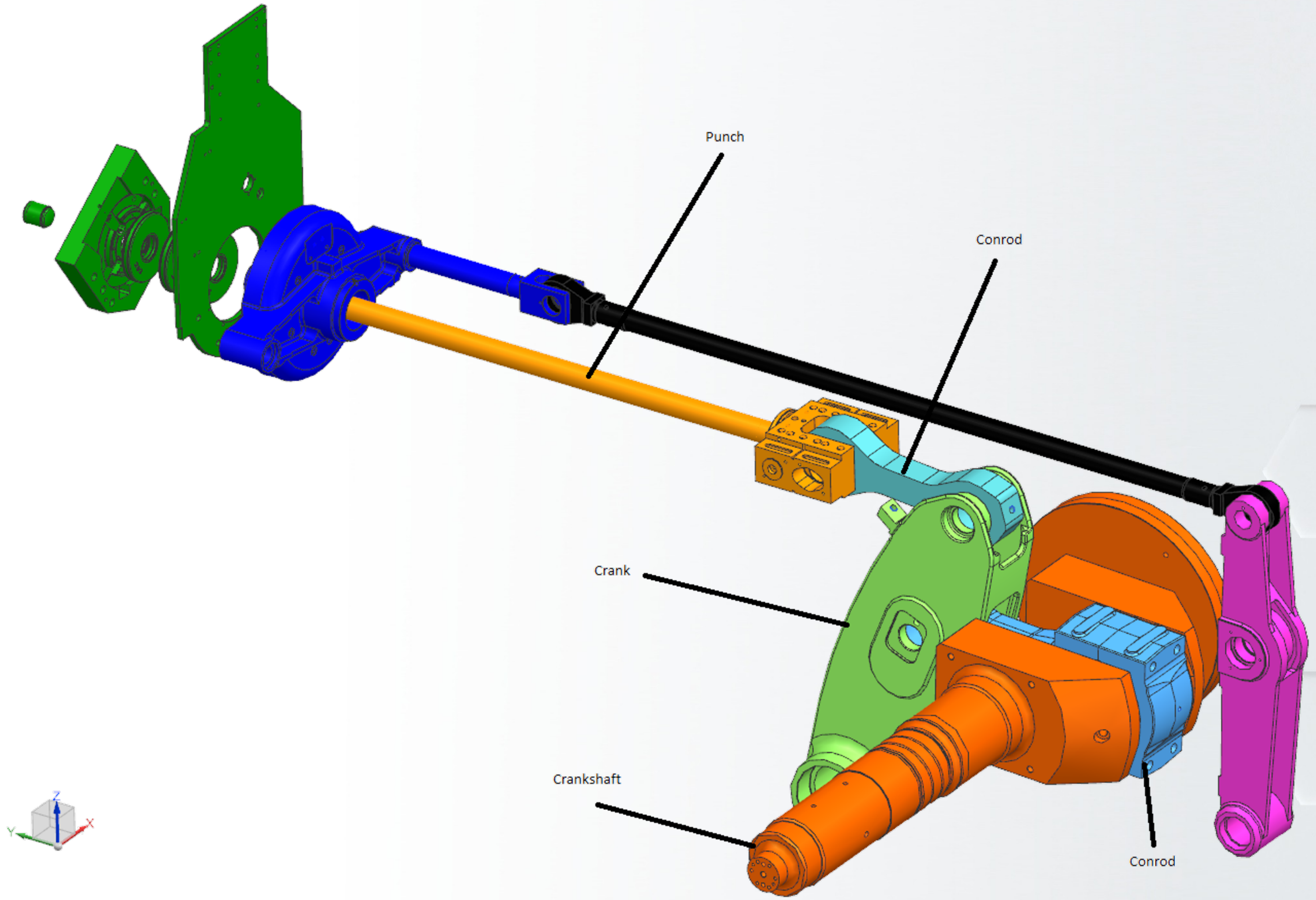
Tear-offs are resolved by removing tooling from the machine and cleaning out scrap cans causing delay in production.

Lost production of 100,000 cans/day

MACHINE TERMINOLOGY OVERVIEW



PUNCH ACTION



TOOLPACK TERMINOLOGY

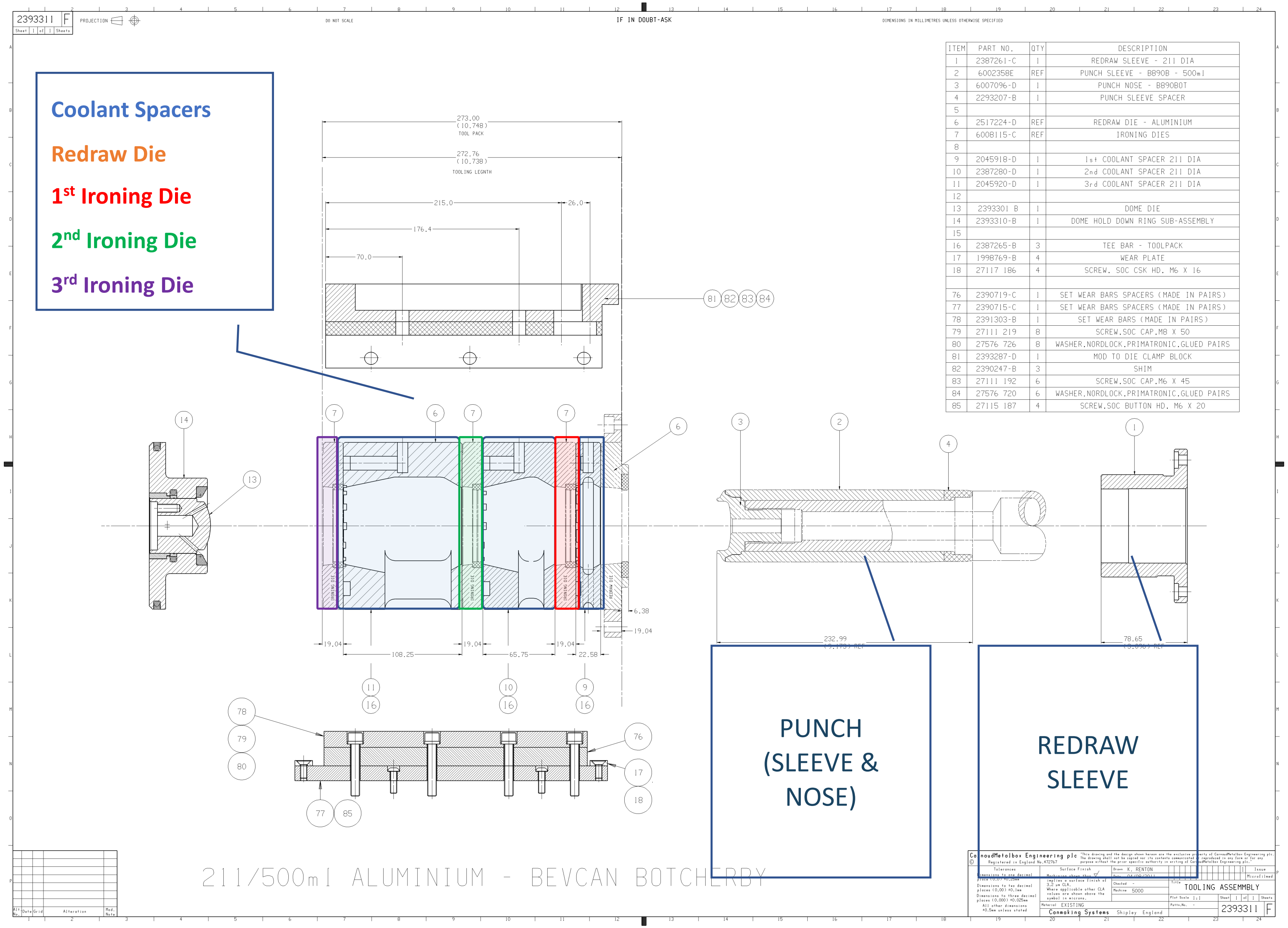
Coolant Spacers

Redraw Die

1st Ironing Die

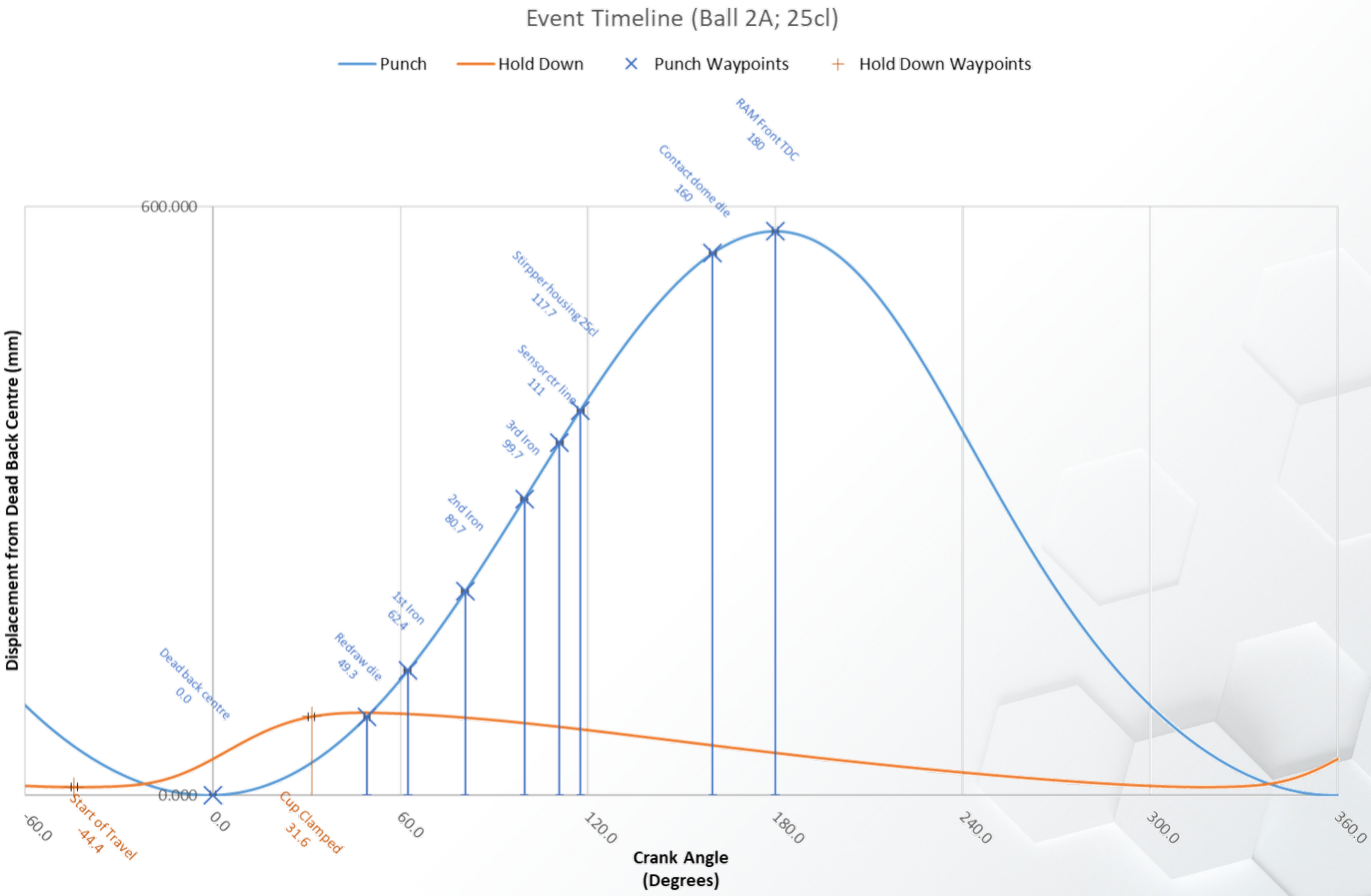
2nd Ironing Die

3rd Ironing Die



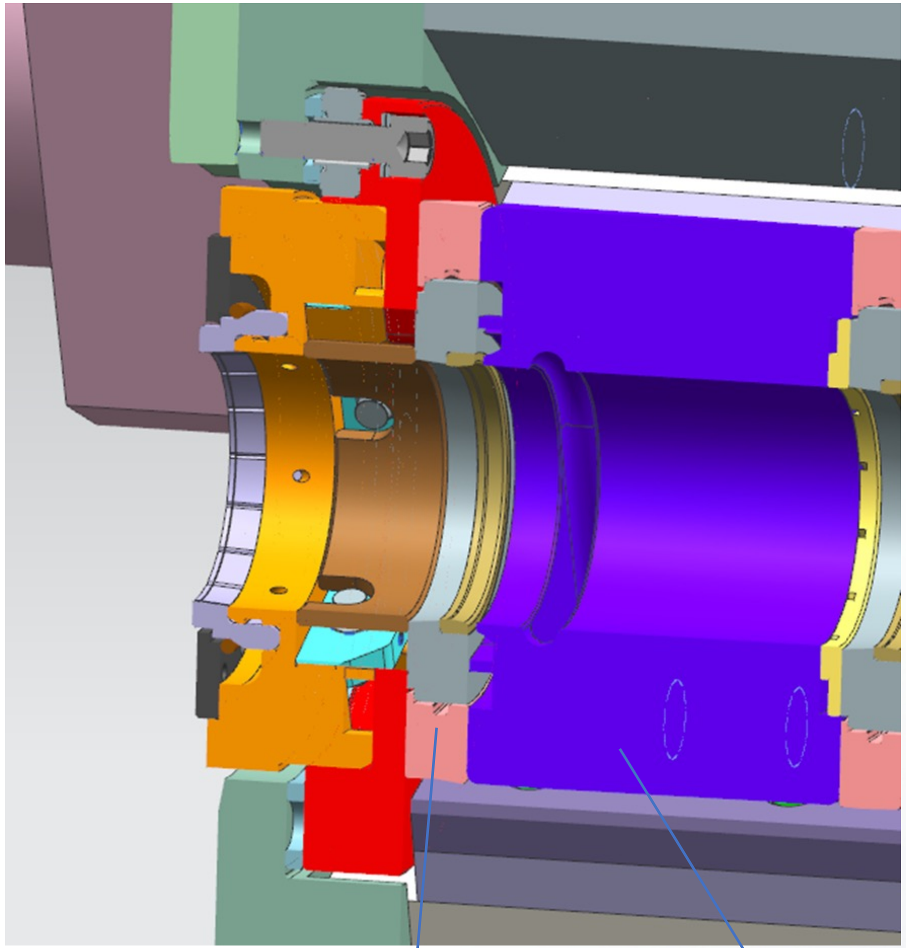
ITEM	PART NO.	QTY	DESCRIPTION
1	2387261-C	1	REDRAW SLEEVE - 211 DIA
2	6002358E	REF	PUNCH SLEEVE - B890B - 500ml
3	6007096-D	1	PUNCH NOSE - B890BOT
4	2293207-B	1	PUNCH SLEEVE SPACER
5			
6	2517224-D	REF	REDRAW DIE - ALUMINIUM
7	6008115-C	REF	IRONING DIES
8			
9	2045918-D	1	1st COOLANT SPACER 211 DIA
10	2387280-D	1	2nd COOLANT SPACER 211 DIA
11	2045920-D	1	3rd COOLANT SPACER 211 DIA
12			
13	2393301 B	1	DOME DIE
14	2393310-B	1	DOME HOLD DOWN RING SUB-ASSEMBLY
15			
16	2387265-B	3	TEE BAR - TOOLPACK
17	1998769-B	4	WEAR PLATE
18	27117 186	4	SCREW. SOC CSK HD. M6 X 16
76	2390719-C	1	SET WEAR BARS SPACERS (MADE IN PAIRS)
77	2390715-C	1	SET WEAR BARS SPACERS (MADE IN PAIRS)
78	2391303-B	1	SET WEAR BARS (MADE IN PAIRS)
79	27111 219	8	SCREW.SOC CAP.M8 X 50
80	27576 726	8	WASHER.NORDLOCK.PRIMATRONIC.GLUED PAIRS
81	2393287-D	1	MOD TO DIE CLAMP BLOCK
82	2390247-B	3	SHIM
83	27111 192	6	SCREW.SOC CAP.M6 X 45
84	27576 720	6	WASHER.NORDLOCK.PRIMATRONIC.GLUED PAIRS
85	27115 187	4	SCREW.SOC BUTTON HD. M6 X 20

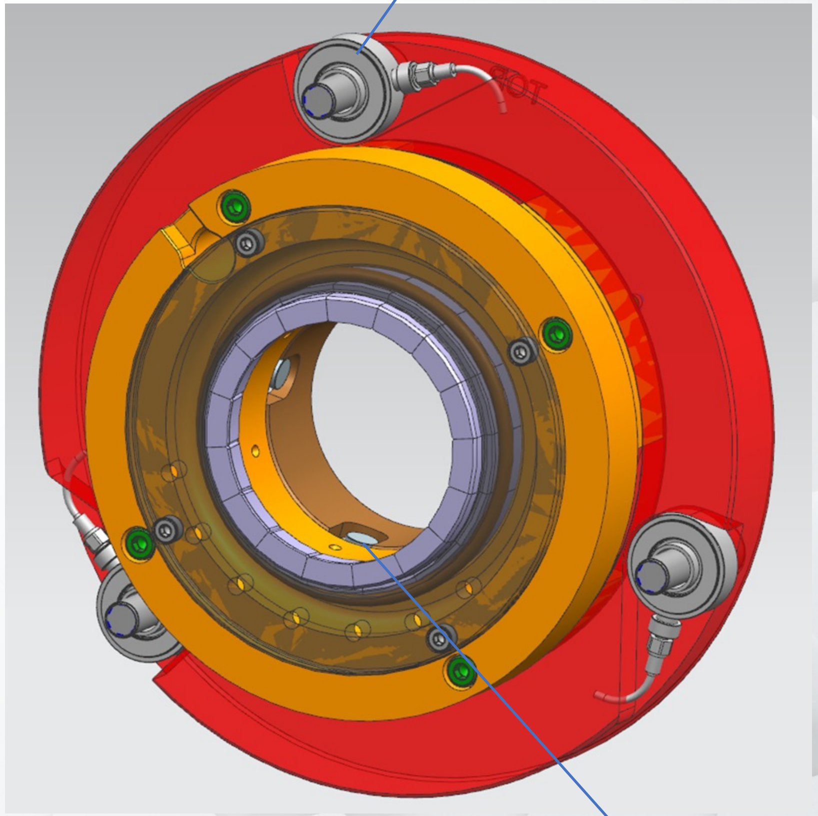
EVENT TIMELINE



SENSOR ARRAY

- Sensor Array is contained within a modified version of the "stripper assembly" – allows for easy retrofit in most CMB Bodymakers.
- During operation, the tool pack (the combination of spacers and Ironing Dies) is held firmly against the stripper assembly wear plate (in red) by means of a pneumatic clamp.
- The proximity of the 3rd iron to the eddy current sensors means we expect to see readings on both sensor types for a period.

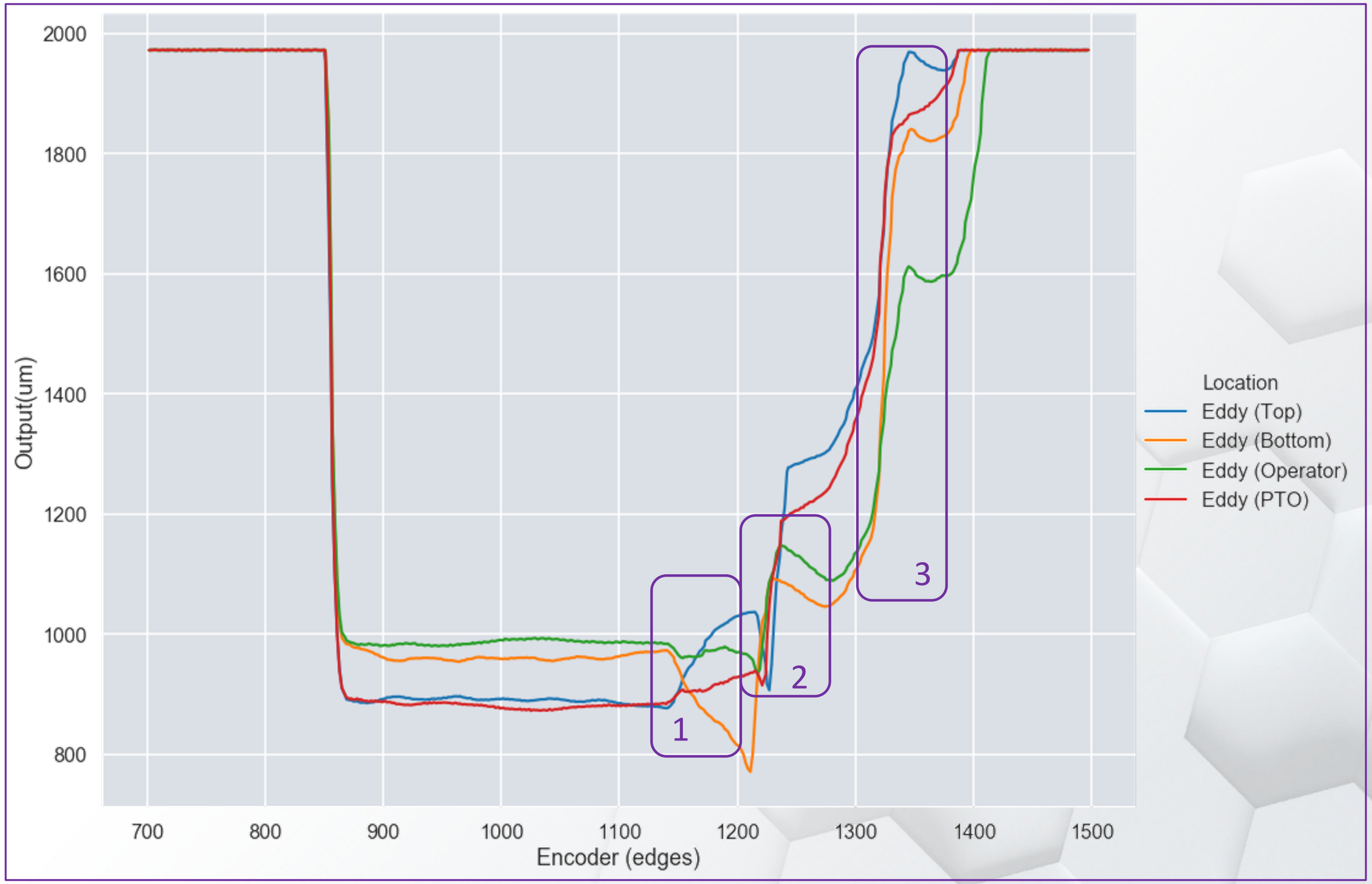




EDDY CURRENT SENSORS

Identifying Features Hypothesis

- Looking at can formation (outbound) section of stroke
- Region 1: Ram deflection after can exits 3rd iron
- Region 2: Can height (edge) detection
- Region 3: End of punch



THE DATA & AI PROBLEM

IoT Data For Prediction: Valuable But Cursed by Abundance and Decentralisation, Creating Cost and Risk

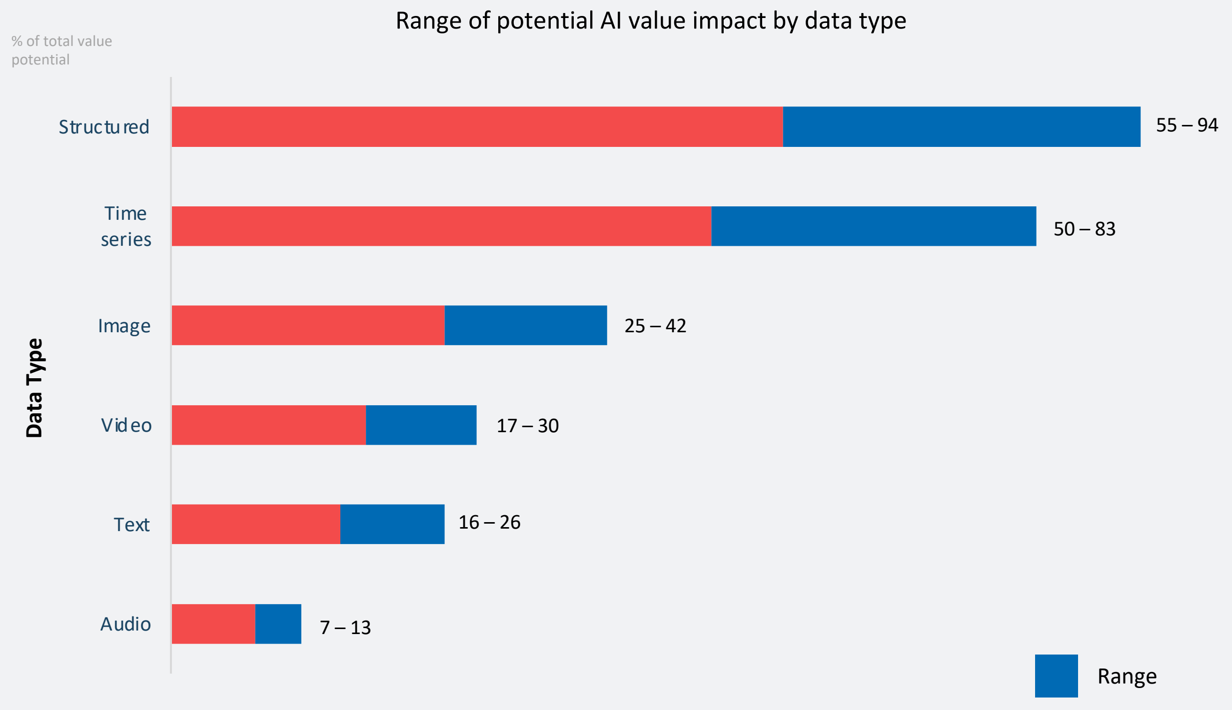
7.6 bn IoT devices in the world - IoT data is the most abundant data type and has the greatest value return from AI.

Traditionally data is centralised for ML/AI.

This creates a problem as IoT data is generated in a decentralised way, meaning that:

- **Complex** and **expensive** to move – Large infrastructure and engineering overhead
- Heavy reliance on **network connectivity**
- **Data privacy** and **security risk** from centralisation

The more data we want to centralise, the greater the opposing forces to doing so.



SOURCE: McKinsey Global Institute analysis

THE DATA & AI PROBLEM

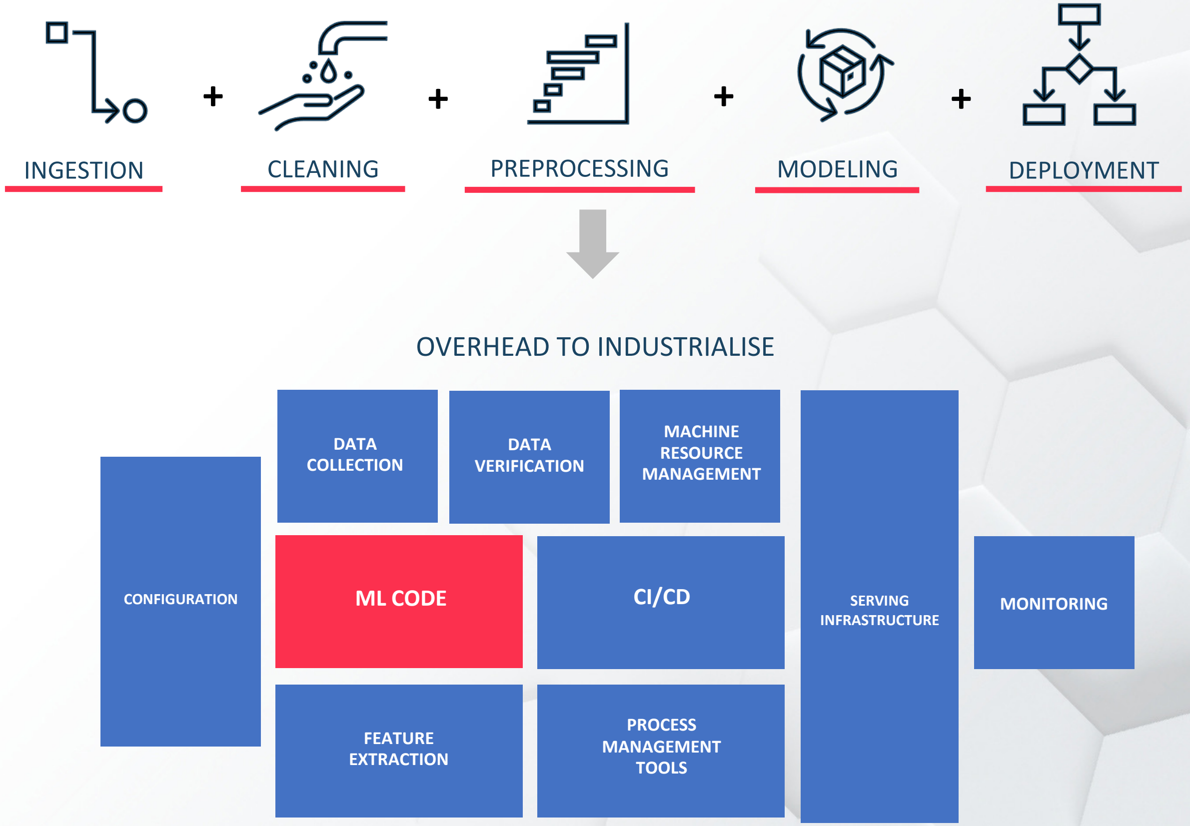
Industrialising AI: Training AI is now easy, but traditional production-grade Machine Learning is hard to scale in the IoT landscape

Reliance on costly streaming of data to and from centralised analytics platforms for prediction for many devices,

Large engineering overhead to automate and scale machine learning systems into production for many devices,

High time and cost of manual development of pipelines,

High cost of and scarcity of in-house AI engineering teams.



A new approach is required for industrialised machine learning in a distributed IoT landscape

> *“75% of manufacturers believe that their greatest barrier to AI is a lack of suitable infrastructure to deploy and manage machine learning in an IoT environment (Google Research, 2021)”*

THE SOLUTION

To Rethink Conventional AI Systems Design for IIoT: A Simple Idea

Don’t move the data to the model… move the model to the data

Scalable infrastructure spanning and connecting the edge and the cloud

\+

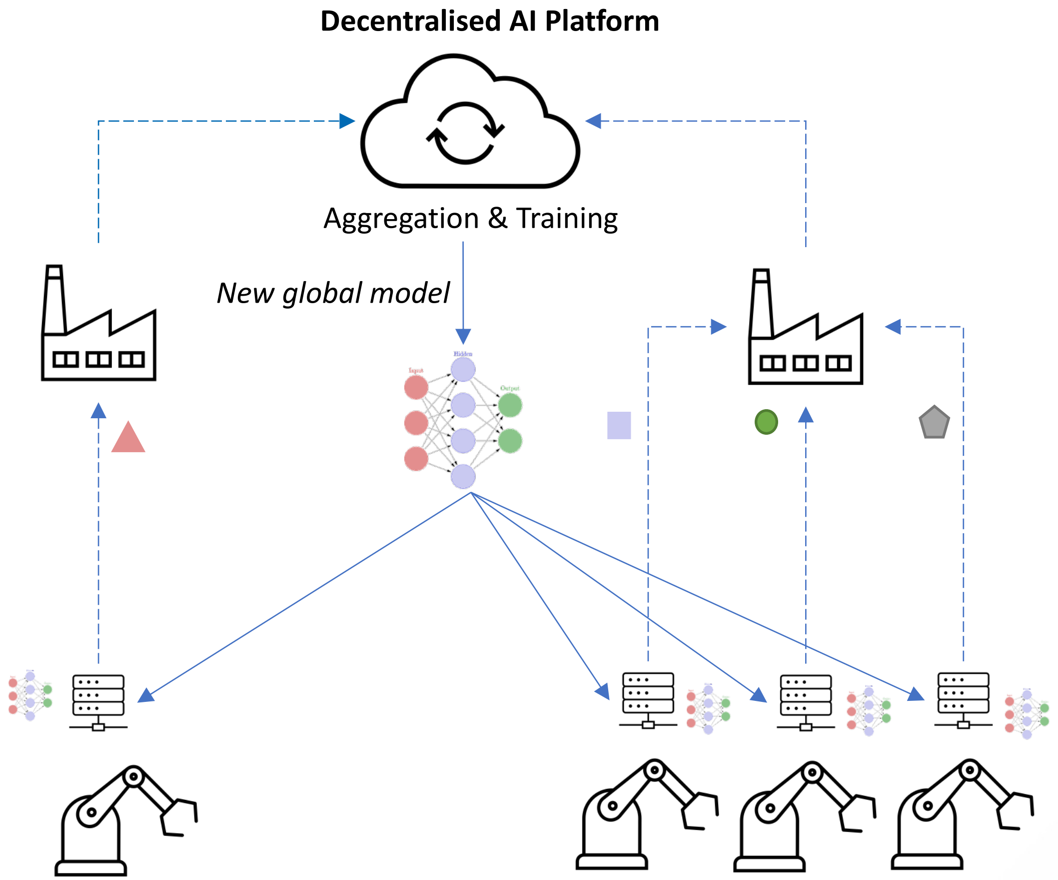
Minimise the movement of data from source devices

\+

Automate the machine learning process to predict machine health

THE SOLUTION

Distributed Learning: A New Paradigm For Decentralised AI



Where possible, rather than move data, we move models to the data

Train models locally at nodes, then aggregate the model weights

Minimizes data transfer and storage costs, maximises privacy and security

Does not require constant network connectivity for inference

Export batch analytics when efficient, run streaming analytics on local compute and network

THE SOLUTION

Cloud connected body makers: Using ML to improve actual operating time

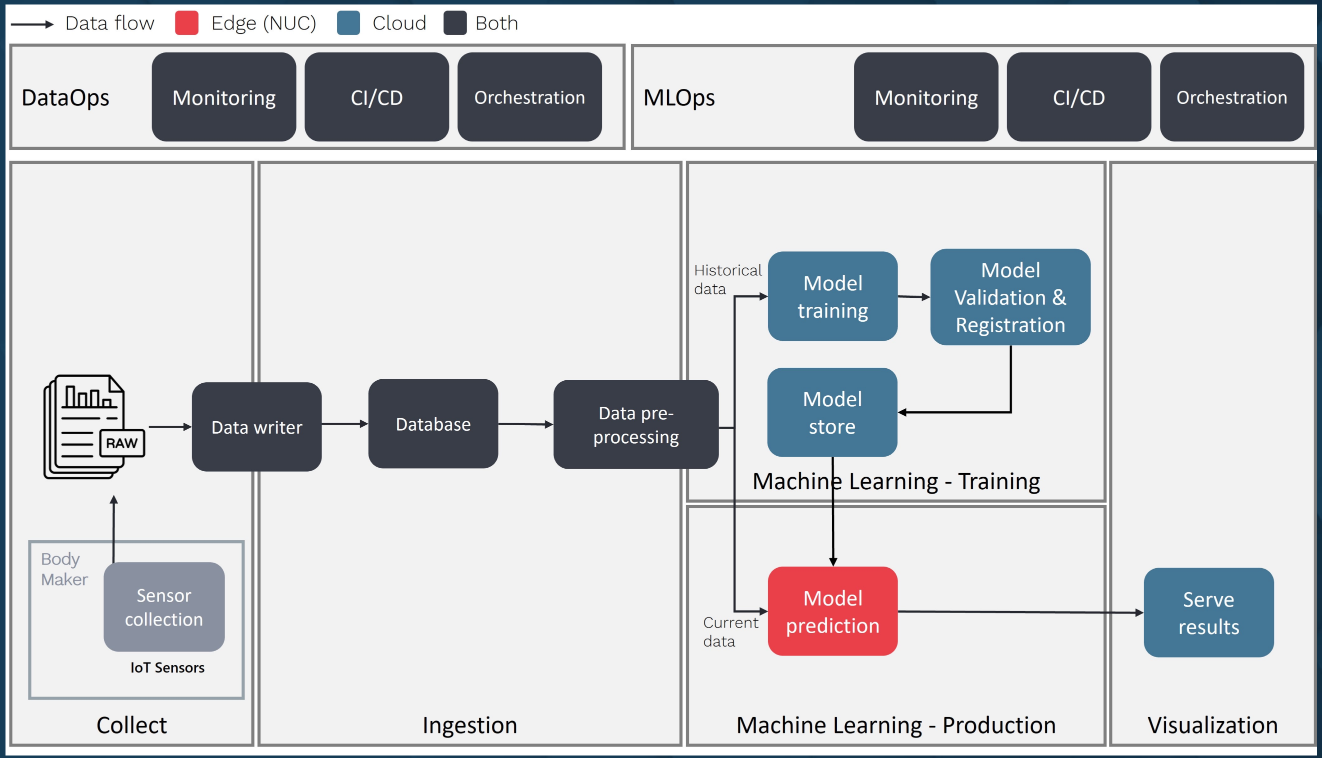
Produce a remaining useful life prediction algorithm that uses the sensor data to predict when a tear-off will happen.

Develop hybrid cloud-edge computing predictive-analytics architecture to analyse bodymaker sensor data and provide feedback to the bodymaker (e.g. realign and optimise speed/forces) and an operator (through a dashboard);

Incorporate sensors with data transfer, storage, and processing capabilities in a miniaturised edge-device sensor and NUC device

Analyse sensor data with machine-learning techniques to identify the precursor signatures of tear-off events;

Deploy trained algorithms to edge devices to monitor and react to bodymaker sensor data in real-time.

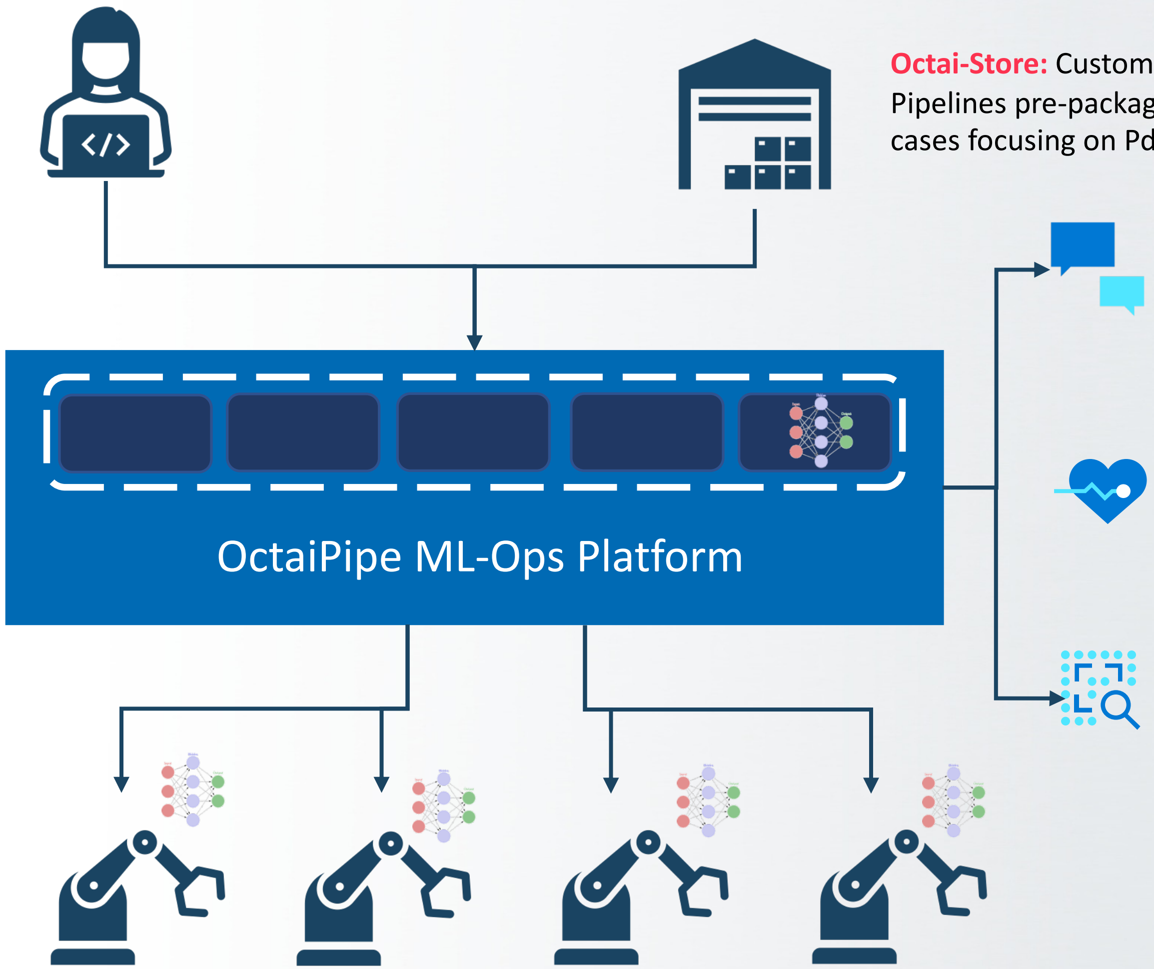


OCTAIPIPE

Octai-Lab: Data scientists can be more efficient in creating and deploying ML pipelines between the edge and cloud on OctaiPipe

Octai-Store: Customers can rapidly deploy ML Pipelines pre-packaged for high value industry use cases focusing on PdM

*Deployments on Octaipipe are more **scalable** and **resilient** due to decentralised design of ML-Ops and advanced decentralised & federated learning.*



Octai-Bot: Conversational AI to troubleshoot issues

Octai-Doc: System health monitoring

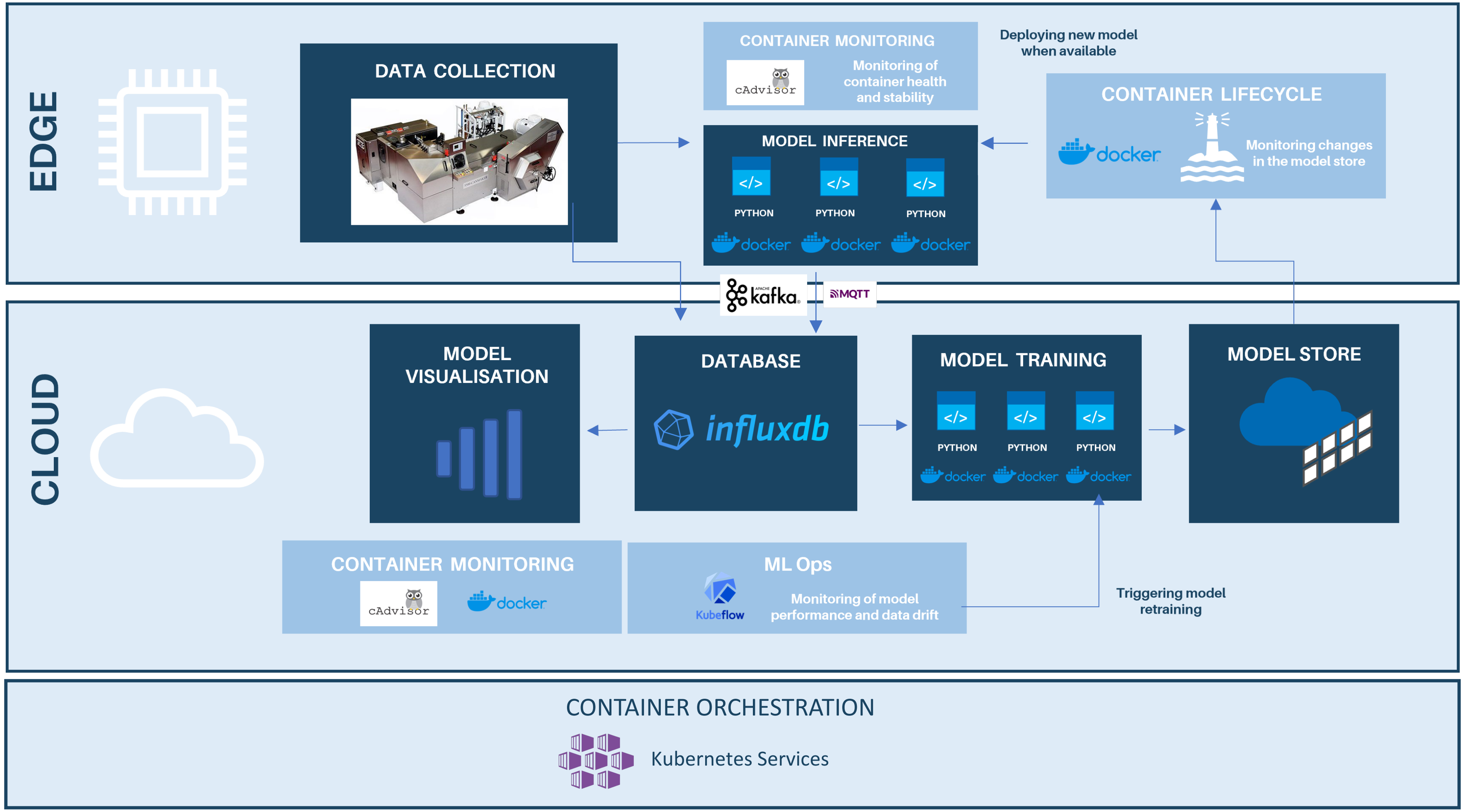
Octai-Sight: Historical and predictive downtime insights and predictions

Zero development costs + **Minimised deployment time and cost** + **Minimised engineering support cost** + **Optimised data transfer, processing, and storage cost** + **Maximised resiliency, privacy, and security** + **Maximised AI scalability** = **Maximised machine uptime & ROI**

OCTAIPIPE

Octaipipe Edge: ML-Ops platform optimised for IoT running across the boundary of Edge and Cloud

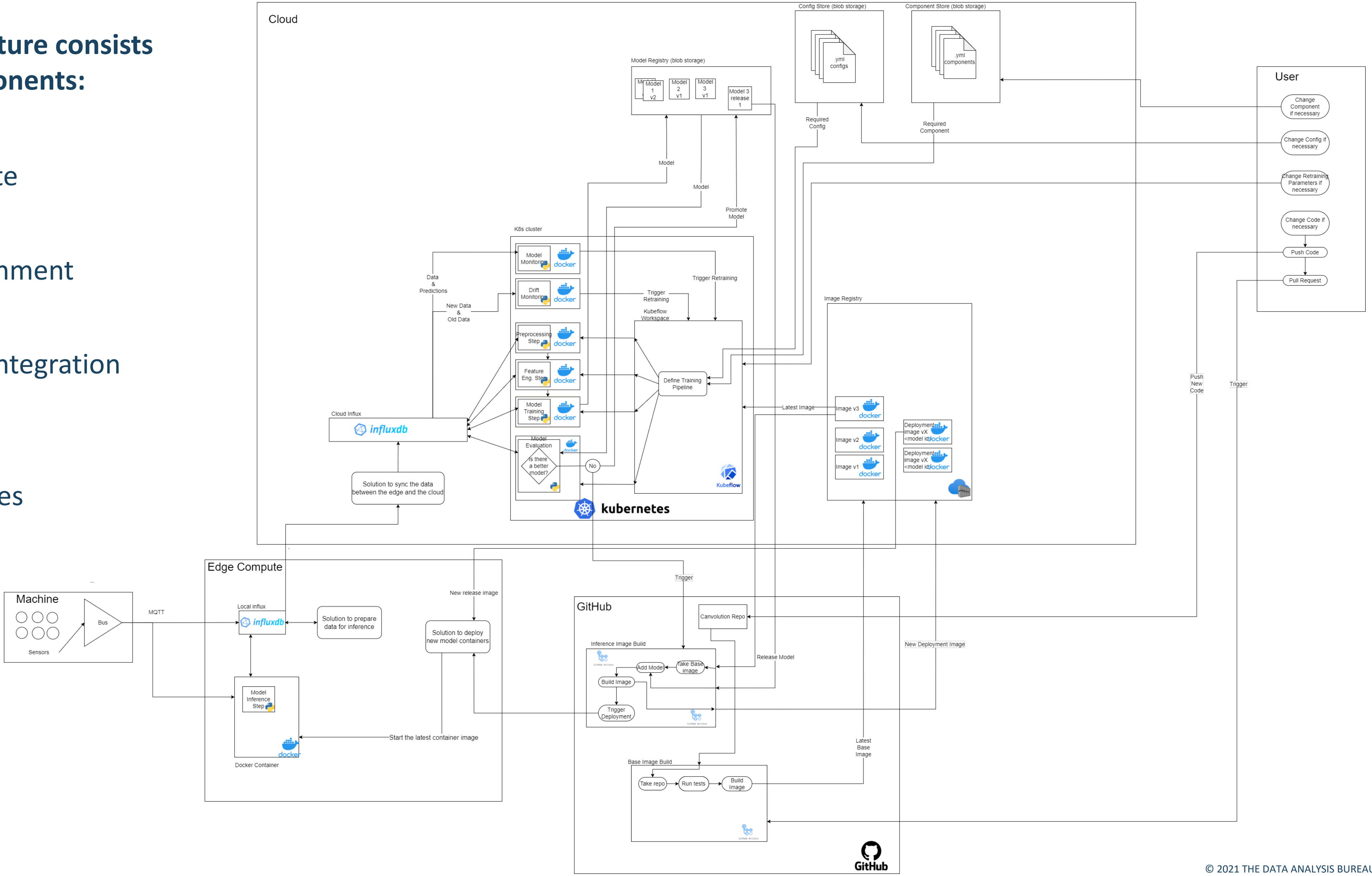
- **Optimised hybrid cloud-edge operation:** Cloud agnostic, optimised ML-Ops shared between edge and cloud
- **Rapid deployment:** Pre-configuration reduces set-up time
- **Network resiliency**: ML inference can run at the edge independently from the cloud. Training can occur independently.
- **Accelerated ML:** Standard ML pipeline template = simple repeatable steps save customer data scientists time to deploy models



OVERALL ARCHITECTURE

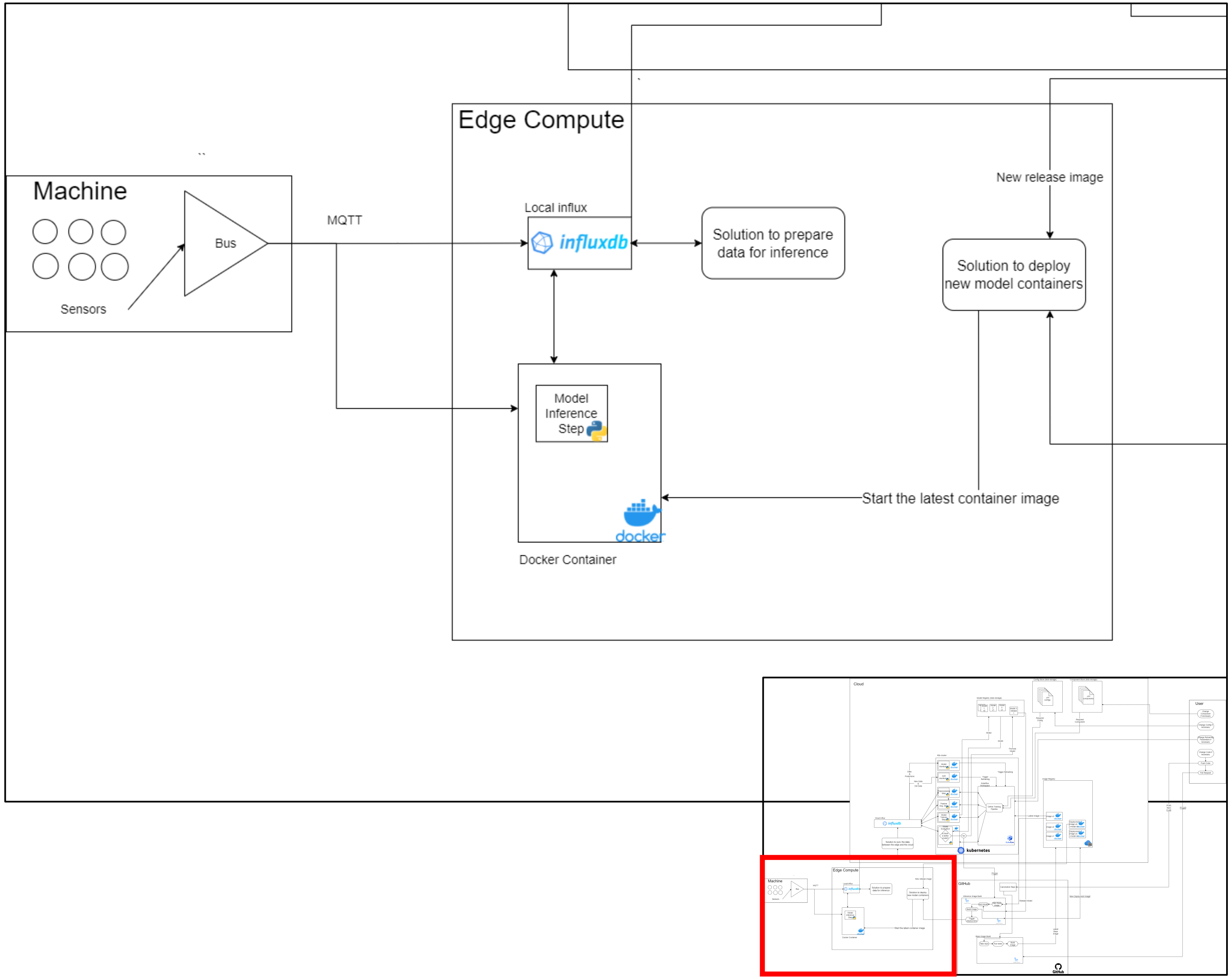
Overall Architecture consists of 4 main components:

- Edge Compute
- Cloud Environment
- Continuous Integration Server
- User interfaces



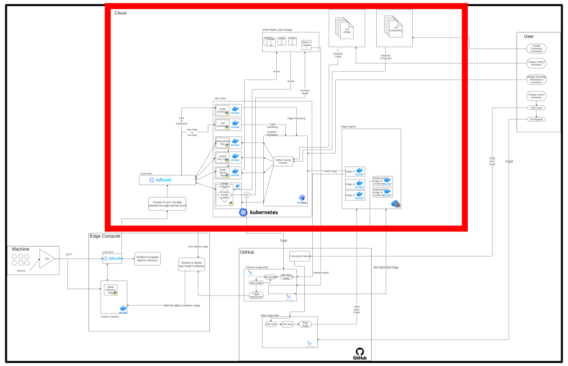
EDGE COMPUTE

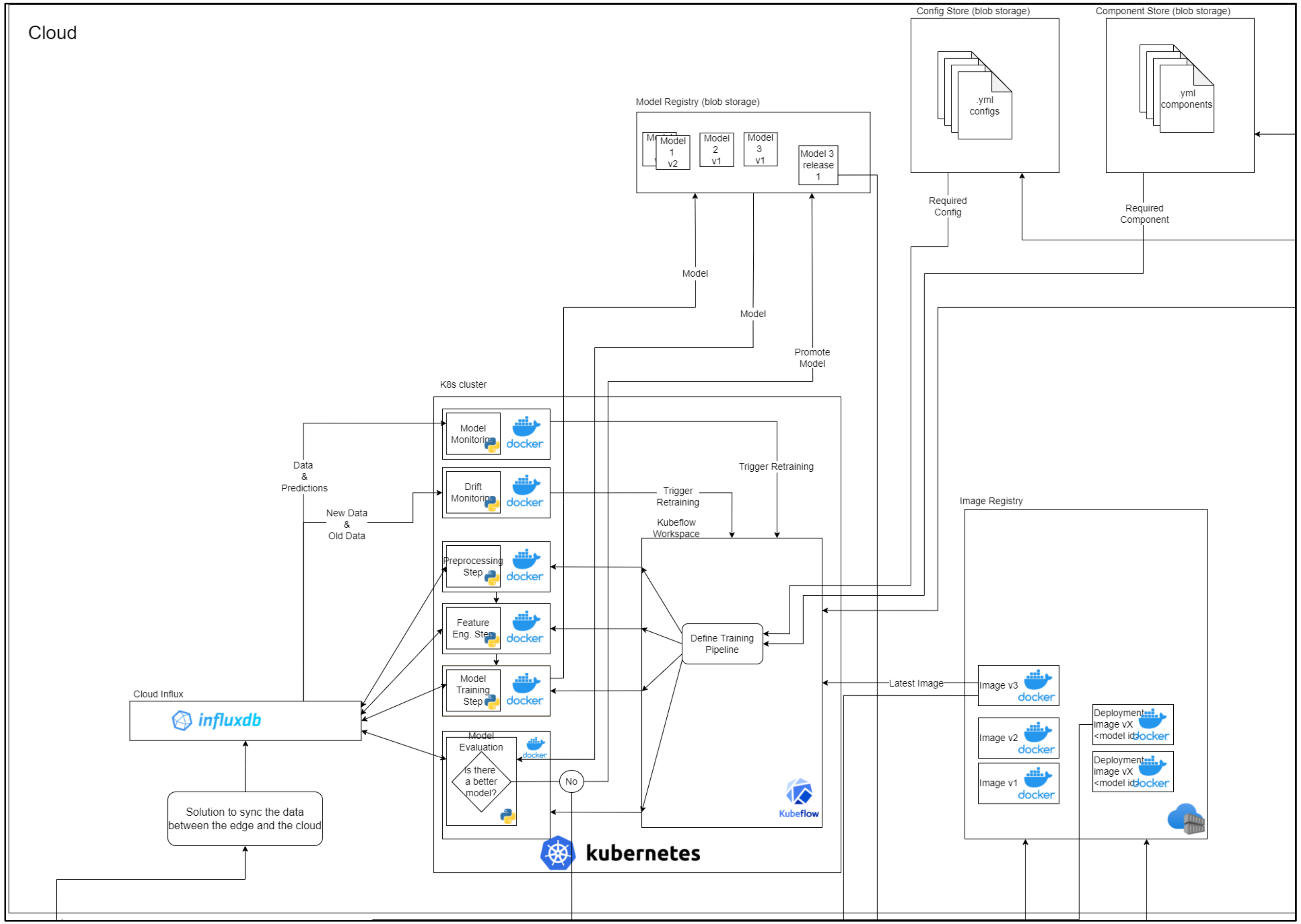
- Data from sensors is analyzed right at the edge
- Database-level transformation is available for speed-optimisation
- Connection to the cloud for data transfer
- Automatic deployment of new model versions



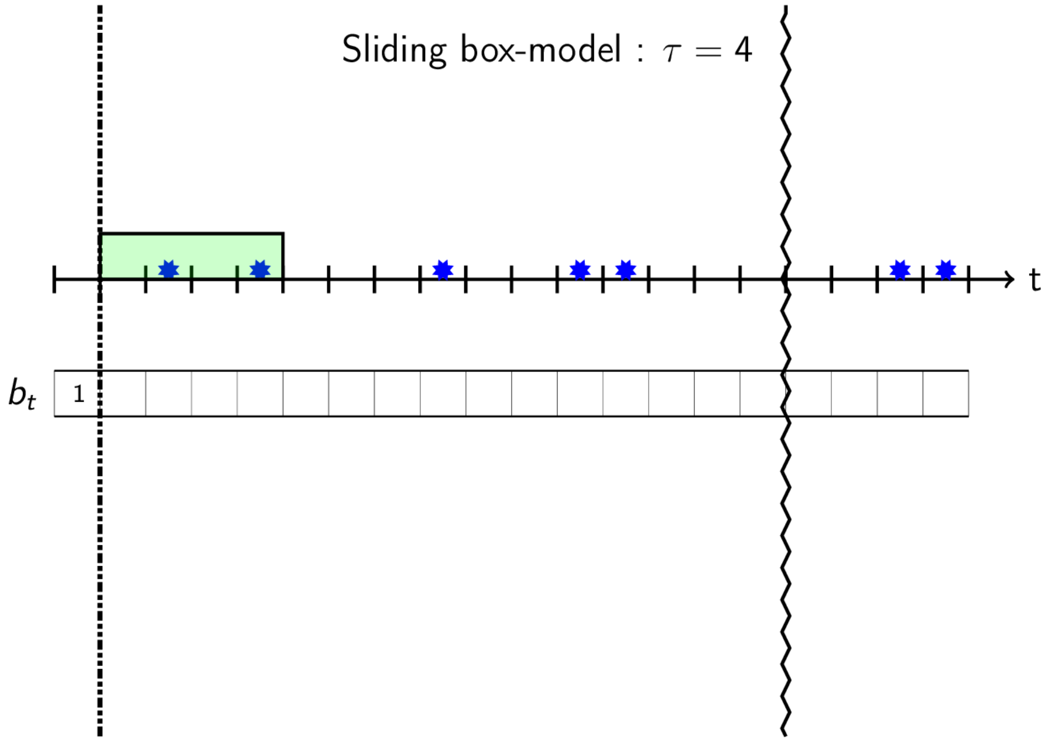
CLOUD COMPUTE & STORAGE

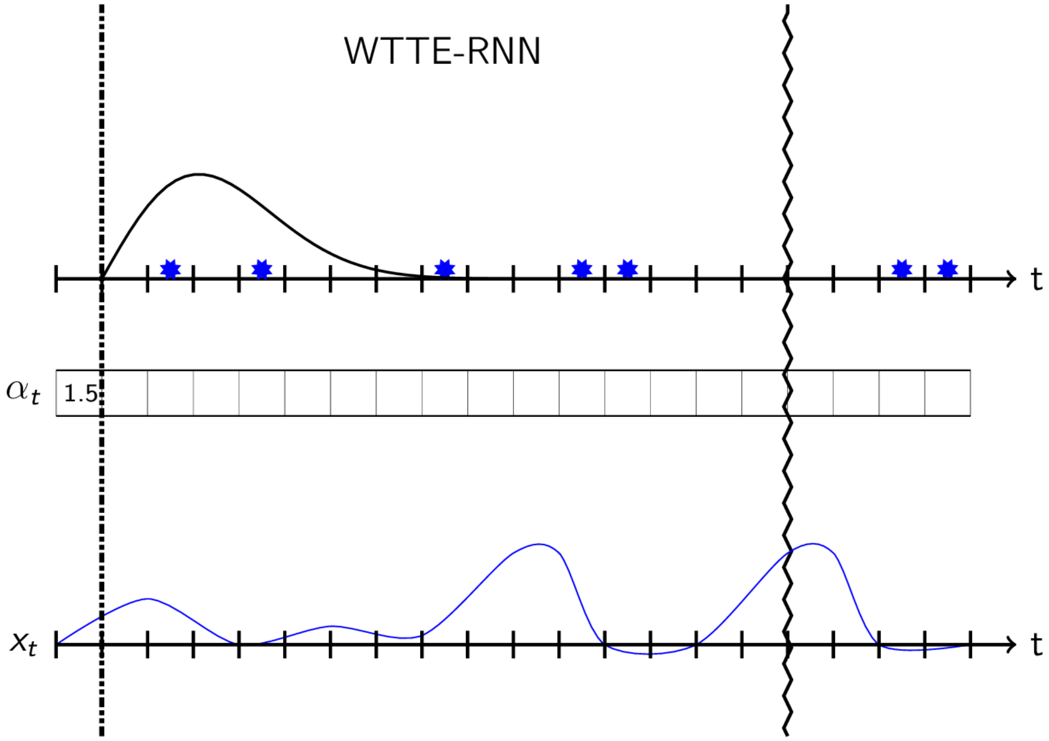
- Model training is carried out from central data store
- Kubeflow and K8s provide computational optimisation and horizontal scalability for the experiments





PREDICTING FUTURE FAILURES: REMAINING USEFUL LIFE WITH OCTAIPIPE





Belimo meets T-DAB, November 11, 2019

REMAINING USEFUL LIFE (RUL) WITH ML

A spectrum of machine learning approaches that is going to be tried out include statistical methods, classical machine learning approaches as well as more advanced deep learning architectures

Survival Analysis

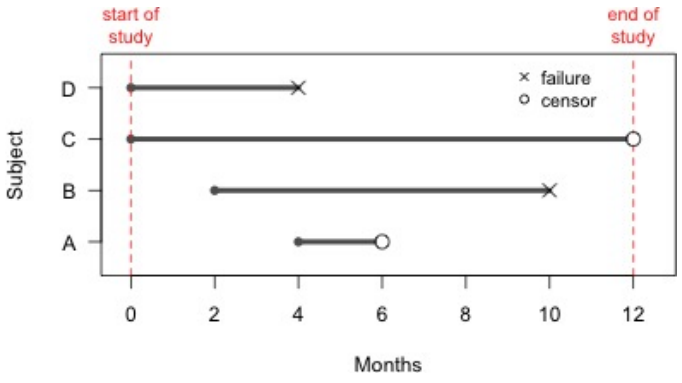
- Cox Regression
- Kaplan-Meyer Estimator

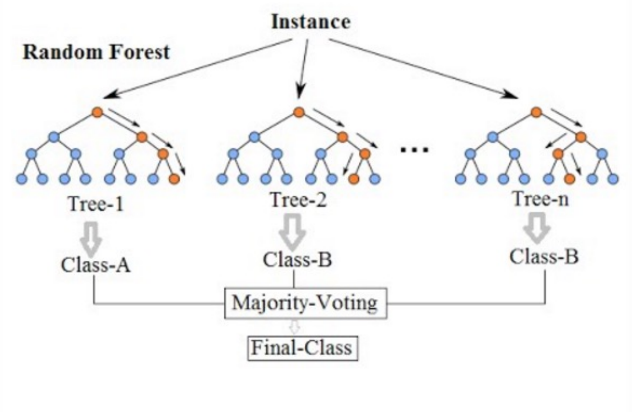
Classical Machine Learning

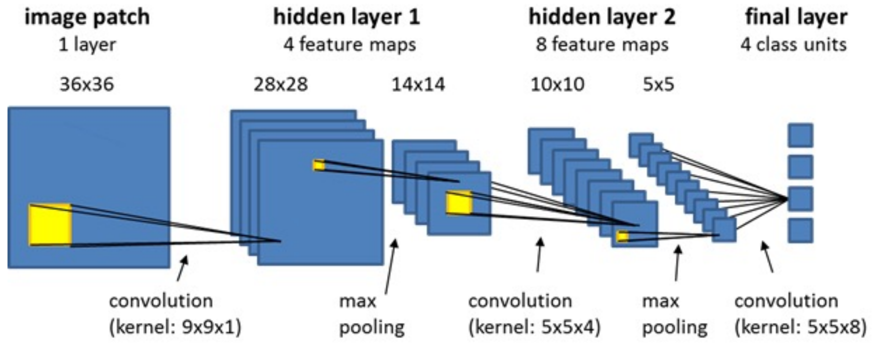
- Decision Tree-based methods
- Support Vector machines

Deep Learning

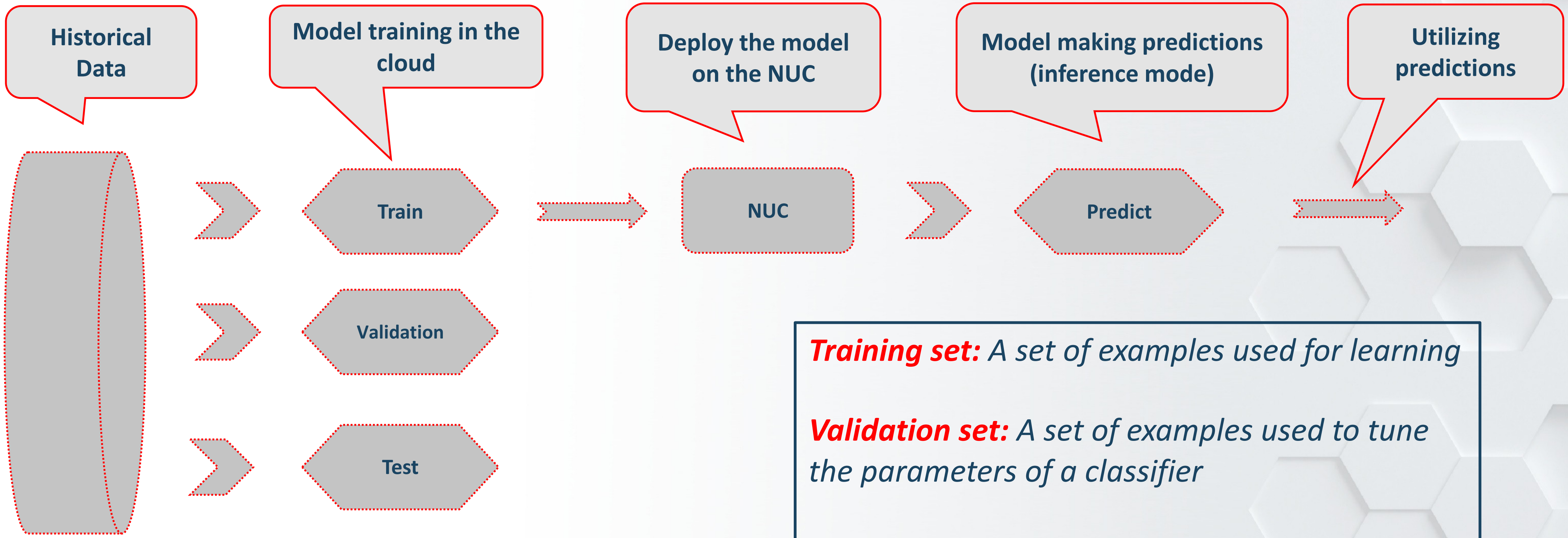
- Recurrent Neural Networks
- Convolutional Neural Networks







ML PIPELINES FOR CLOUD CONNECTED BODY MAKERS



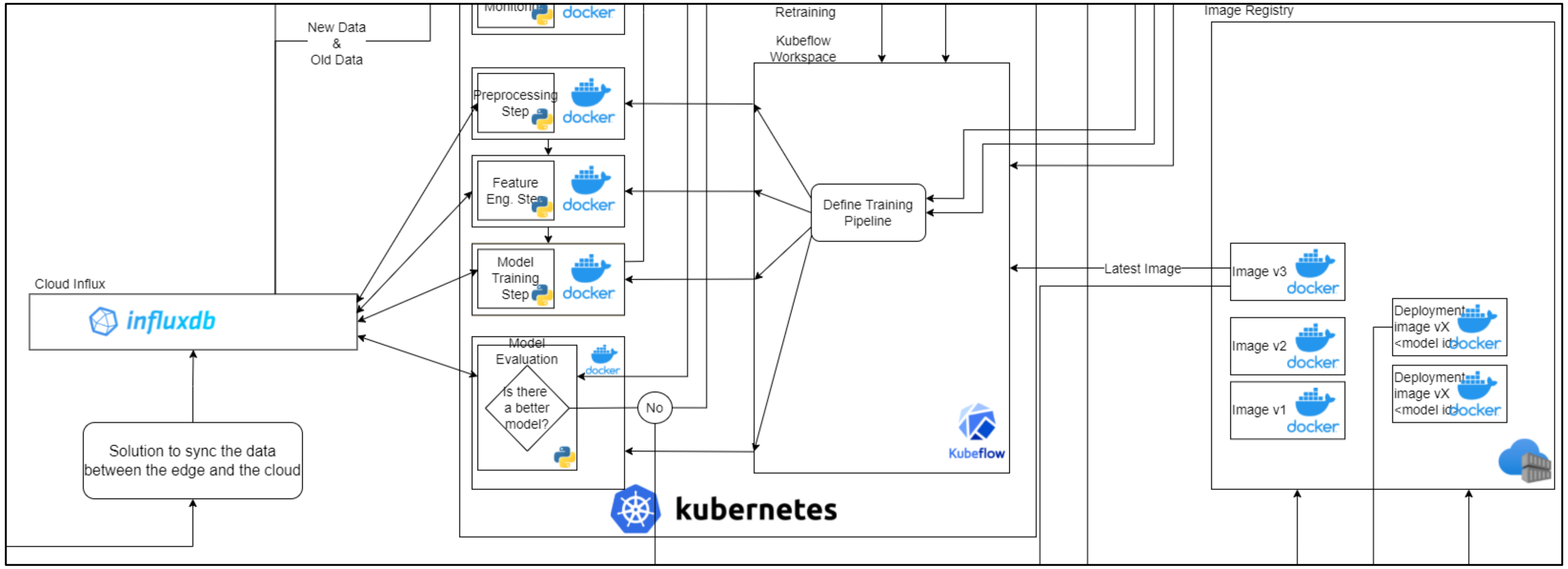
Training set: *A set of examples used for learning*

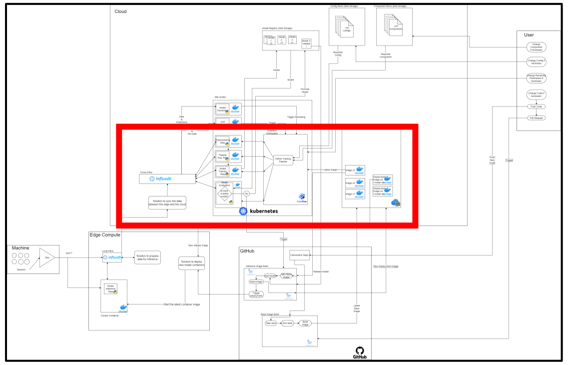
Validation set: *A set of examples used to tune the parameters of a classifier*

Test set: *A set of examples used only to assess the performance of a fully-specified classifier.*

AUTO-ML MODEL TRAINING PIPELINES

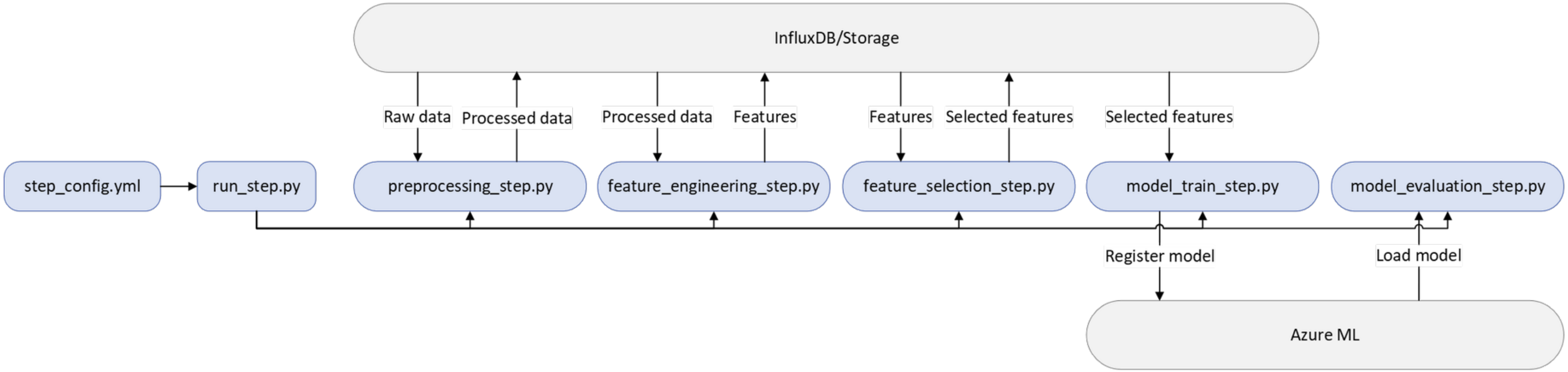
- Training pipelines are pre-defined defined using Kubeflow
- Each pipeline step runs independently
- After training, the model is evaluated and deployed if required





AUTO-ML PIPELINES

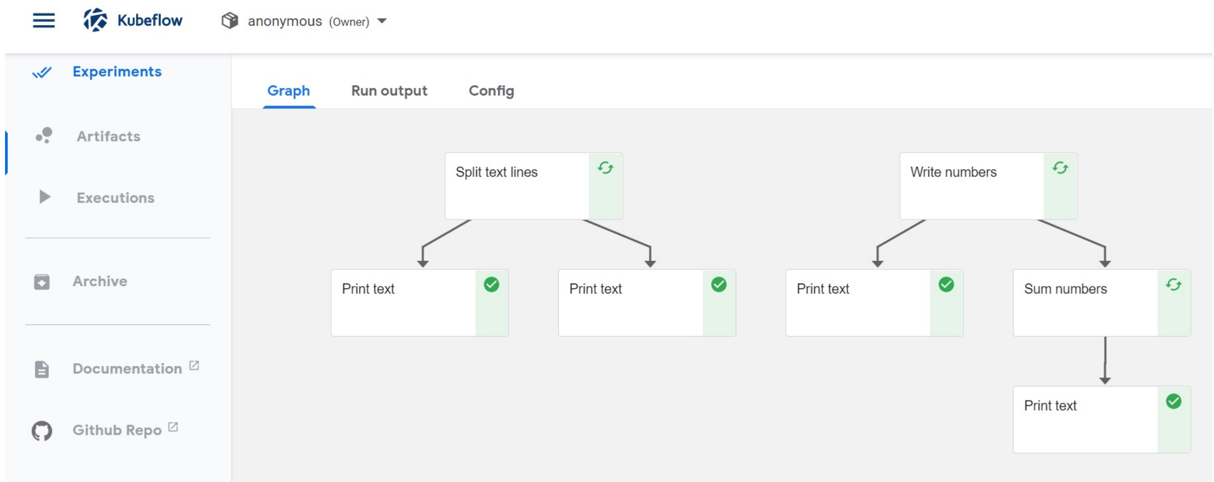
Use Case Optimised, Pre-fabricated Machine Learning Pipelines

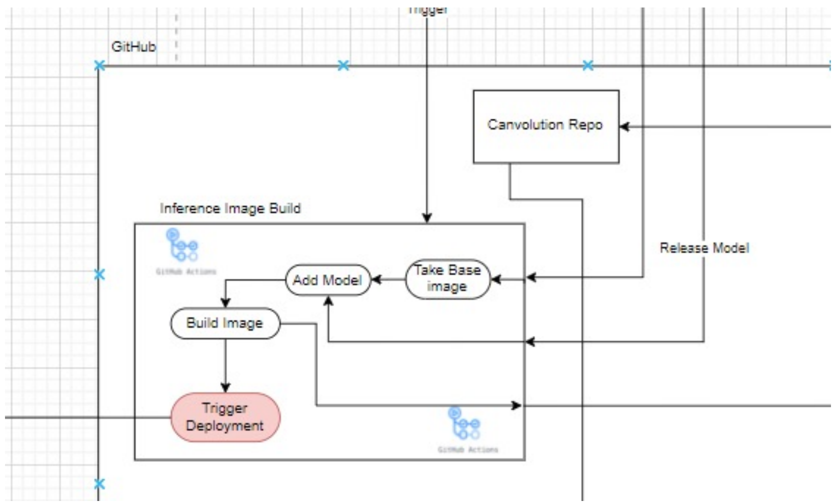


- The manual data science process can be automated
- **'Pipelines'** - a series of automated functions applied to the data
- Carefully designed for specific, known industry use cases
- Example: data pre-processing for RUL estimation with pressure and temperature data will be different to that for Quality Assessment using images and anomaly detection
- Training and inference pipelines can be wrapped into a single program and deployed to the edge to run automatically

AUTO-ML TRAINING & INFERENCE PIPELINE ORCHESTRATION

- For quick experimentation, simple python functions can be run in Kubeflow
- Created a process for constructing training pipelines
- Inference image build is triggered after model evaluation
- The required model is loaded directly into the docker image





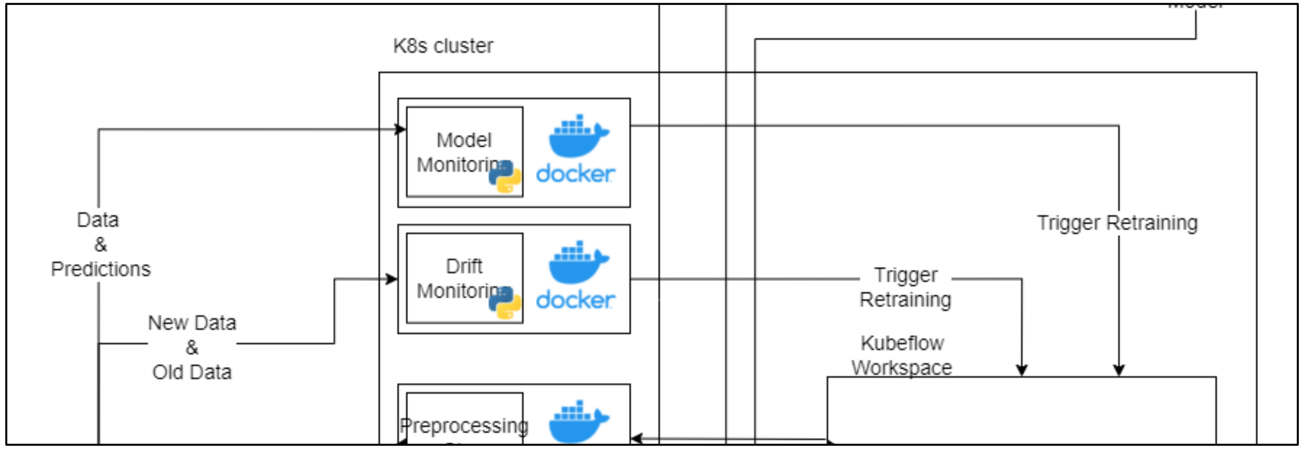
OUTCOMES

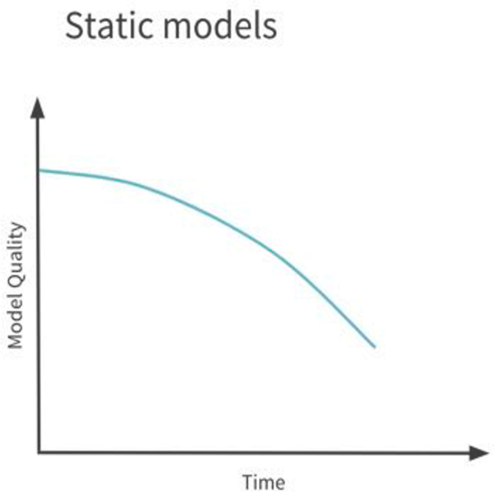
- High accuracy = Errors of only 10 – 15 operating cycles (mean 250 cycles = 6% error in estimation)
- Auto-ML pipeline and Federated ML performance as good or better than benchmark state-of-the-art
- Performance dependent on model type, aggregation algorithm, and data split
- Required no or minimal movement of sensor data beyond that needed for analytics task.
- Only of pre-packed pipeline, evaluation metrics, and model weights

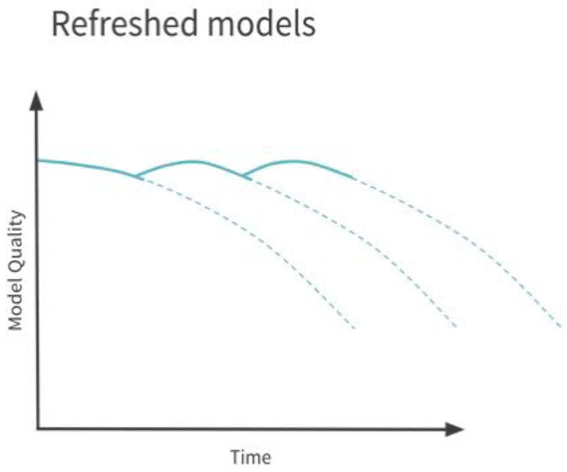
Models	Test RMSE	Test MAE
External Benchmark Models		
Random Forest [48]	12.01	not performed
Deep LSTM [50]	18.43	not performed
Hybrid SVR/LSTM [51]	19.11	not performed
Hybrid CNN/LSTM [57]	13.85	9.44
Fed. MLP (FedAvg/FedProx) [18]	20.90 - 23.92[1]	16.41 - 18.70[1]
Selected Centralised Models		
Kaplan-Meier	31.86	19.74
Cox PH	36.95	30.82
Random Forest	16.63	12.24
Neural Network	20.80	15.42
Selected Federated Model		
Fed. GBDT (imbalanced data split)	23.36	18.07
Fed. GBDT (balanced data split)	19.15	13.81
Fed. NN (imbalanced data split)	17.37	13.21
Fed. NN (balanced data split)	20.74	14.82

ML-OPS MONITORING

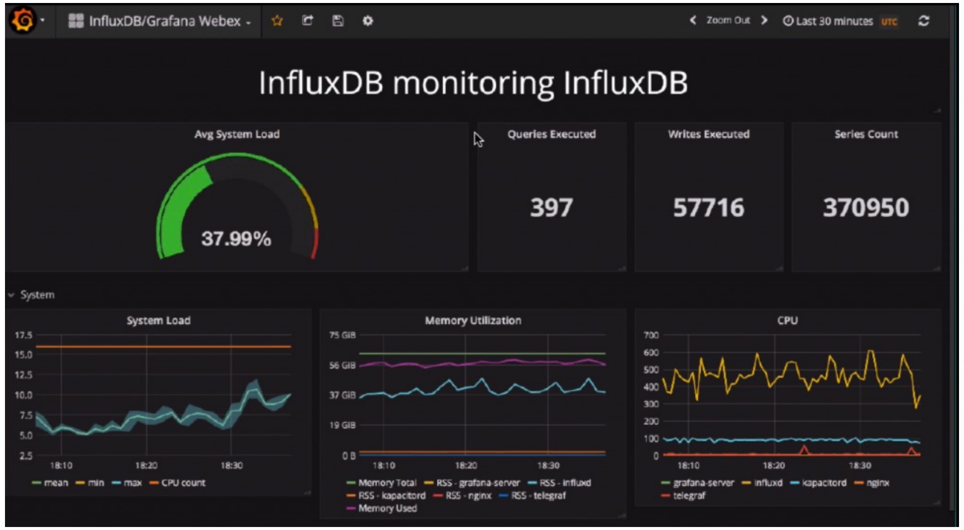
- Key challenges are:
  - Data drift
  - Model drift
  - Model versioning and CI/CD
- Machine learning powered AI systems are like human co-workers; they need training, appraising, and retraining to stay relevant
- Model performance and data are monitored from additional AI services on the cluster
- Retraining is triggered or recommended if a decline in performance or drift in data distributions is identified is identified by monitoring algorithms
- Human-in-the-loop components provide extra level of robustness







FRONT-END: INSIGHTS AND MONITORING

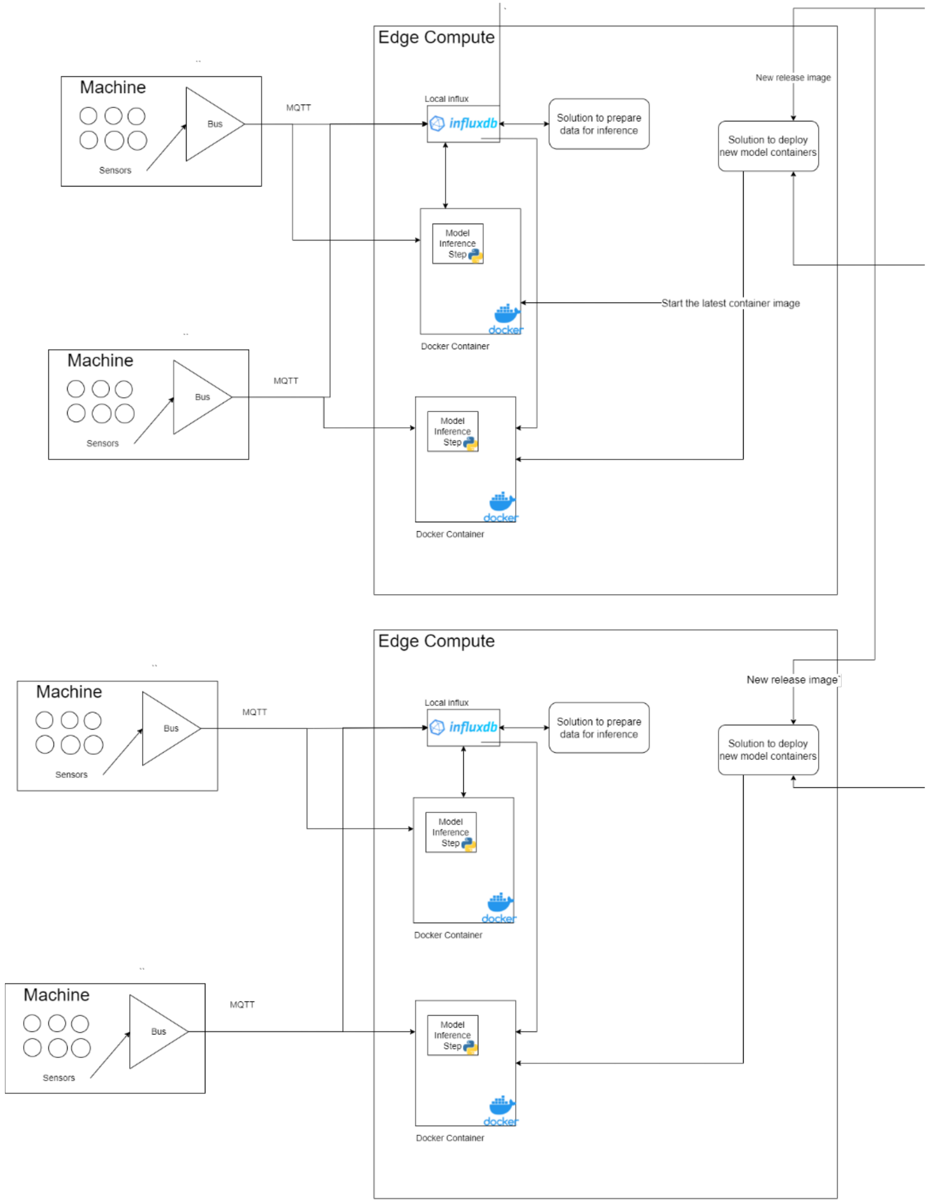


- A dashboard for system health is used to monitor for cluster resources utilization, health of docker containers, databases, and success of the CI activities
- Insights dashboard shows all the necessary data summarized to the level of user: high-level overview for managers, factory-floor insights for engineers, etc

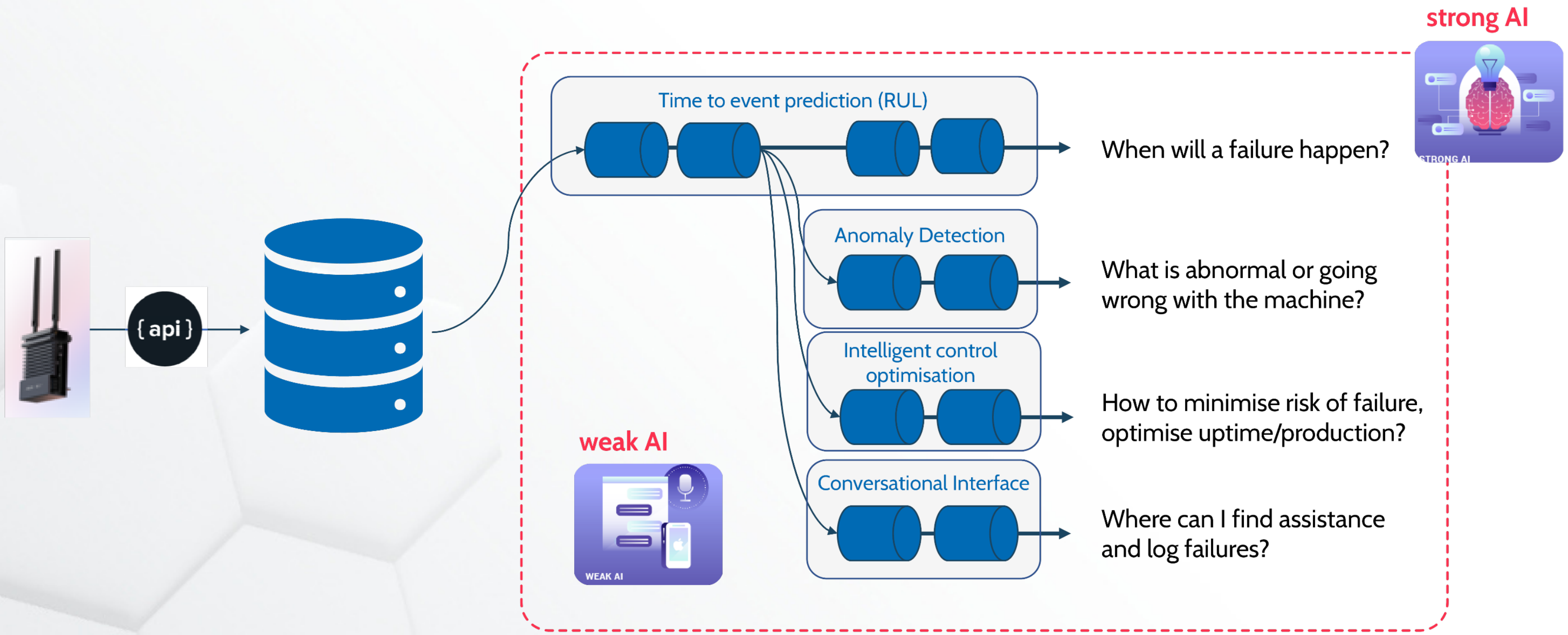
SCALING TO MULTIPLE DEVICES

- Multiple machines can be connected to the same edge device which will be running multiple model containers

- For higher number of machines, extra edge compute can be added and connected to the same cloud and CI systems as before



SCALING TO MULTIPLE USE CASES



- Create Strong AI-like systems from weak AI components
- Efficiency conveyed by foundational upstream building blocks being reused by downstream pipelines
- Containerised design with common data models, standardised I/O , and connectors
- Allows ML pipelines with complementary use cases to be connected and scaled

BENEFITS

BUSINESS

1. Prediction of tear-off events leads to reduced downtime through stoppage prevention
2. Analysis of tear-off risk allows for machine control optimisation to minimize stoppage risk
3. Anomaly detection allows intelligently targeted maintenance, reducing required downtime
4. Down time forecasting allows for better planning of maintenance, improving overall uptime and productivity
5. Better understanding of the physical process causing downtime leading to best practice improvement and standardisation

TECHNOLOGY

1. OctaiPipe reduces the engineering overhead (time and cost) to build infrastructure and deploy ML intelligence – currently less than one month to deploy
2. OctaiPipe optimises data transfer and compute cost for analytics and machine learning intelligence
3. OctaiPipe in-built AI powered ML-Ops monitoring improves system resilience and reduces the number of required engineers
4. OctaiPipe reduces reliance on network connectivity to central cloud

Questions?