



Orid Dynamics

**Reference Architecture Technical Paper** 



## **Edge Computing for Retail**

Reference Architecture using PowerEdge XE2420 Server

September 2020

#### Abstract

This document outlines Reference Architecture that utilizes capabilities of the Dell EMC PowerEdge XE2420 for handling significant AI workloads related to running an edge-computing powered service for retail needs in a secure environment. The scope includes implementing four use cases of typical AI usage in retail.

## Contents

Executive summary	3
Retail Edge Analytics System overview	4
Use Case 1: eCommerce-style analytics for Brick & Mortar stores	5
Store key performance indicators and forecasts	6
In-store traffic counting	6
Store manager dashboard	6
Use Case 2: Advanced in-store behavior analytics	6
In-store traffic measurement in multiple locations inside the store	7
Analytical dashboard	7
Use Case 3: Inventory decision support tool for store managers	7
Store-level demand forecast	8
Comparison with the corporate forecast and alerting	9
Use Case 4: Anomaly detection in server metrics	10
HTM for real-time anomaly detection	10
Retail Edge Analytics System architecture on Dell EMC PowerEdge	11
Implementation	11
Use Cases 1 + 2: Recognizing and counting traffic	11
Use Case 3: Smart inventory management	15
Use Case 4: Detecting anomalies on the server	17
Reference architecture model and configuration	19
Server Configuration Used	19
Deployment process	19
Reference Architecture	20
Benefits of using Dell EMC PowerEdge XE2420 as the edge server for retail	20
Use case architecture	21
Conclusion	22

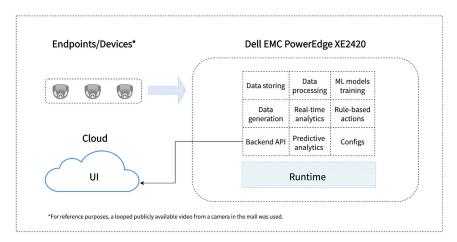
3

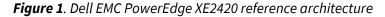
#### Executive summary

According to KMPG's *"Living in an AI World 2020 Report*," AI will prove to bring the local-level execution to the next level. Moreover, AI, supported by unsupervised machine learning (ML), can capture and track customer experience in real time and without bothering the visitors. Advanced computing models such as artificial intelligence (AI) already impact the business significantly. To compete against digital-native competitors, physical retailers have to adjust both their business model and day-to-day operations.

Edge computing drastically reduces the load by providing computing and analytical resources locally, in a secure and autonomous environment. However, many retailers find it hard to build and maintain systems needed to provide the infrastructure for AI initiatives. Centralized solutions, despite their numerous advantages, introduce organizational hurdles. So, because of data latencies or losses due to compression as well as high operational costs, these are not the best choice for the retail industry.

To demonstrate how Dell EMC PowerEdge XE2420 server handles significant AI workloads, the reference architecture below was used to simulate the computational needs of retail and related IT infrastructure for a particular store.





This paper is designed for IT Systems Architects to handle the IT needs for their Edge specifically looking into the retail environment.

The reference architecture shows transforming large volumes of in-store data into AI-powered retail intelligence and addresses the challenges of four use cases:

- eCommerce-style analytics for Brick & Mortar stores for making data-driven decisions based on the key sales and traffic metrics.
- Advanced in-store behavior analytics for tracking store visitors patterns and dwelling areas.
- Inventory decision support tool to keep inventory optimized against actual demand and plan for the stock replenishing.
- Anomaly detection for server metrics for providing central IT support or upgrades before any downtime happens.

#### Retail Edge Analytics System overview

Al and Machine learning (ML) based platforms for consumer understanding can help retailers make focused investments to enhance customer experience as well as positively impact the bottom line. Finding areas for improving the experience starts from *understanding the customers and their reaction when they happen without bothering them*.

Dell Retail Edge Analytics system reference architecture shows the benefits of using Dell EMC PowerEdge XE2420 server as an environment for creating flexible moduled Point of Sales (POS) system that:

- handles large volumes of end-to-end processing data;
- analyzes key metrics in real time;
- provides insights locally;
- is easy to upgrade and maintain;
- runs securely and is not dependent on the internet connection;
- detects possible issues in advance to prevent system downtime;
- has low adoption friction.

The system runs locally from a web browser. The 4 use cases chosen to demonstrate the Dell EMC PowerEdge XE2420 server capabilities are described in the corresponding sections below. Each use case represents a reference analytical solution to gauge customer experience (CX) retail intelligence, inventory management, and unsupervised smart alerts for any anomalies on the server.

This section guides through the business aspect of the system. To see the technical details and implementation walkthrough, see *Retail Edge Analytics System Architecture on Dell EMC PowerEdge*.

#### Use Case 1: eCommerce-style analytics for Brick & Mortar stores

User persona	Store manager
Business value	Tracking key metrics of the store operation in real time, comparing them with historical information on an intuitive analytics dashboard. Making data-driven decisions based on forecasts of sales and traffic insights.

Using Dell EMC PowerEdge XE2420 server as the Server platform for getting data from the in-store channels to analytical insights on a simple dashboard. For the reference, a publicly available video from a CCTV camera was used. The streamed video was ongoingly processed, and the details on each visitor stored into a database. Large sets of historical data were also generated to show the capabilities and speed of data processing by Dell EMC PowerEdge XE2420 server.



Figure 2. eCommerce-style analytics for Brick & Mortar stores analytical dashboard.

#### Store key performance indicators and forecasts

- Aggregated sales and expected sales forecasts
- Distribution of sales by gender: aggregated sales history and AI-powered forecasts.
- Average transaction statistics and predictions made by complex algorithms.
- Historical trends of in-store traffic and forecasts using ML models.

#### In-store traffic counting

- Visitors profiling: using AI-powered people recognition, distinguishing visitors by their gender.
- Live traffic analysis: counting the visitors in real time, then aggregating the numbers into daily statistics and computing forecasts for a better queue management and staff distribution.
- **Peak traffic insights**: using visitors history to predict the day and time with the highest amount of visitors in the upcoming week.

#### Store manager dashboard

- Maximizing performance with advanced analytics: all data available on the dashboard, allowing store managers to see trends, identify areas for opportunity, and react to exceptions.
- Driving efficiency: Aggregated stats for current week and comparison with the last year.
- **Forecasts for justified decisions**: make more informed decisions by analyzing future trends computed with state-of-the-art approach in time-series forecasting.

#### Use Case 2: Advanced in-store behavior analytics

User persona	Store manager
Business value	Mapping in-store journeys of customers to direct optimal product placements, monitoring how much time they spent in zones and what regions drive the most attention. Identifying the best and the worst performing categories of products.

As a logical continuation of store analytics use case, the reference architecture utilizes data gathered from CCTV cameras to determine the regions of high interest and plot them on a heatmap. All data transformation is performed on the Dell EMC PowerEdge XE2420 server to show its ability to handle significant workloads.

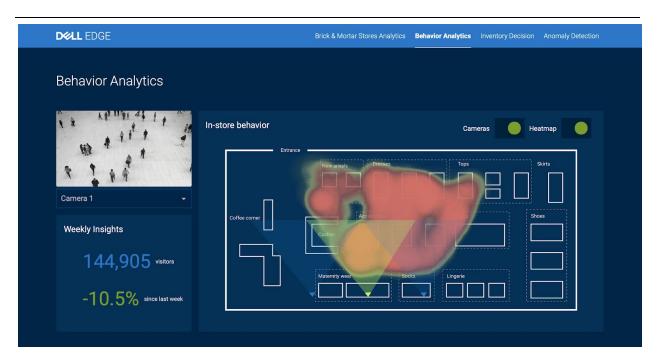


Figure 3. Advanced in-store behavior analytics UI.

In-store traffic measurement in multiple locations inside the store

- Visual discovery using people counting and recognition: depicting regions of interest on a heatmap mapped on a store layout.
- AI/ML to analyze in-store customer behavior: using video from CCTV cameras to detect how visitors dwell in the store.
- Identifying ways to optimize customer experience: adjusting the store layout to better meet customer and business needs.

Analytical dashboard

- **Customer path monitoring**: tracking visitor dynamics on a weekly insights block to determine if assortment needs to be optimized.
- **Capturing customer behavior in real time**: gaining customer insights by analyzing their real-time movements and patterns without bothering them.

#### Use Case 3: Inventory decision support tool for store managers

**User persona** Store manager, regional managers

# Business valueDetecting discrepancies between the planned and real-time POS data, managing<br/>inventory mismatch and optimizing it. Identifying and resolving issues with<br/>under- and over-stocking, generating stock forecasts.

This use case demonstrates the local adjustments for inventory plans for a particular store compared to the central corporate system forecasts. Doing the what-if analysis helps retailers track and adjust stocks against the demand curve as well as plan seasonal challenges ahead. The maximum acceptable difference between demand volume and corporate system's forecast can be set by the user.

For reference purposes, the simulated data along with a linear model trained on this data was used.

Store-level demand forecast

- **Optimizing store assortment**: comparing forecasts (past and next week), historical and current actual demand and stock level, which is aggregated for all stock keeping units (SKU) as time-series plots.
- Identifying items that require attention: determining whether the deviation between demand and forecast exceeds the specified acceptable value (threshold) to plan the store stock replenishments or detect unexpected sales fluctuations.



Figure 4. Inventory decision support tool UI.

Comparison with the corporate forecast and alerting

- **Data at a glance**: all useful data plotted on a bar chart for stock level, line charts for sales and corporate forecast.
- **Browsing the state of stock**: by moving a weekly frame, managers can explore the available data and review the alerts for SKUs, for which the deviation between demand and forecast is greater than the threshold value.
- **Configurable notifications**: managers can specify the acceptable deviation threshold that triggers alert if the root-mean-square deviation (RMSE) between forecast and actual demand is greater than the threshold for a particular unit.

#### Use Case 4: Anomaly detection in server metrics

User persona	IT administrator
Business value	Ensuring the seamless performance of a server by receiving alerts in case of any potential issues. Ensuring maintenance and troubleshooting of the server itself with a least possible amount of human intervention and downtime that may negatively affect daily store operations.

In the reference architecture, volumes of data were simulated and modeled after <u>integrated Dell</u> <u>Remote Access Controller (iDRAC)</u> telemetry. Since it contains preconfigured metric endpoints for anomaly detection and alerting mechanisms, they were used to train Hierarchical Temporal Memory (HTM) models for each metric. Using the models, the reference architecture provides an unsupervised Al mechanism to detect anomalies and adjust the configuration for specifics of the monitored server workload.

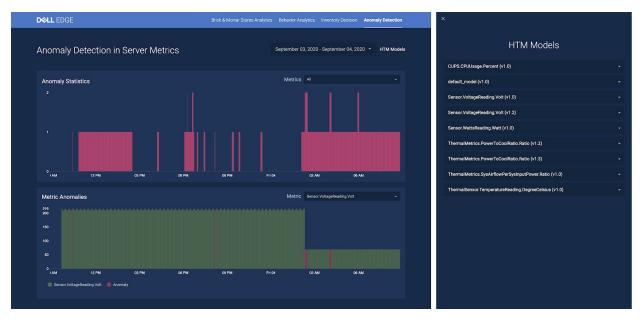


Figure 5. Anomaly detection in server metrics dashboard and configuration menu.

#### HTM for real-time anomaly detection

The term "anomaly" here refers to an outlier — a value that is outside of specified thresholds. Additionally, the term 'anomaly' can be applied to specific time points or longer time periods where a continuously abnormal behavior of time series values is observed. Currently, the solution defines an anomaly as any behaviour of a time series value that is different from the typical signal pattern within a set time frame. This AI driven use case implements a flexible approach based on normal state and behaviour patterns extraction, but does not rely on purely statistical methods. Therefore, it can catch not only suddenly occurring outliers but can also reveal changes in the distribution of very noisy data.

With continuous feed of data from iDRAC Telemetry that includes several of the hardware components of the server like HBAs, COMMs, CPU, Memory, Drives Power etc., an ML model is used to understand the pattern and this trained data set is utilized to detect anomalies which can then be used to predict hardware failures and hence proactive action can be taken by central IT to minimize the downtime.

The anomaly detection reference architecture has been designed and implemented with the following features and constraints in place:

- **Real-time data ingestion and real-time anomaly detection**: simulated data was used for the reference architecture, but it completely mimics a live data stream from a working iDRAC.
- Models real-time training and adoption: historical data is aggregated and served as data sources for future analysis and real-time anomaly detection.
- New metrics onboarding for anomaly detection in runtime: collects system and application metrics through metric collection agents and processes every new time point and metric value.
- Intelligent alerting and decision making: alert configurations can be created for one or more anomaly graphs. Each email contains information about a group of anomalies (abnormal metrics). The quantity and density of anomalies per alert are determined by business rules.
- **Visual assessment**: time series are used for visualising distribution and abnormal behaviour for both the current time as well as times in the past.

### Retail Edge Analytics System architecture on Dell EMC PowerEdge

#### Implementation

To make the application easy to show and access without any additional installations, the user interface was created to be accessible through a web browser. The intention was to simulate the workload for analyzing camera streams and significant volumes of statistical data on a local server. In this manner, the capacity of Dell EMC PowerEdge XE2420 to ensure seamless experience for AI-powered retail intelligence systems was researched.

**Note**: Dell Retail Reference Architecture currently uses the data generated with mathematical models and retail industry knowledge. Its primary purpose is to show how Dell EMC PowerEdge XE2420 meets the needs of retailers in handling large volumes of data and ongoing AI/ML workloads.

#### Use Cases 1 + 2: Recognizing and counting traffic

For simulating one of the pressing areas in which AI optimizes the retail operations, a use case of counting and segmenting visitors as well as their in-store behavior was chosen. Another part of the use case is related to analyzing the sales and traffic trends and utilizing complex algorithms to make predictions.

Key services and technologies

Use Case 1:	Python, MySQL, YOLOv3&YOLOv4, DeepSORT, Pytorch, OpenCV(CV2), Facebook
	Prophet, Apache Airflow, classic analytical geometry approaches (point-line
	distance, half-plane detection, etc.)

**Use Case 2:** Python, MySQL, YOLOv3&YOLOv4, DeepSORT, Pytorch, OpenCV(CV2), classic analytical geometry approaches (homography projection)

#### Real-time video streaming

In the reference architecture, a looped video of real people walking in the mall was used as a source (*see Figure 5*).

#### Image recognition

Powerful ML models, YoloV3 and YoloV4, were used for object detection in images. For tracking visitors, the Deep Sort algorithm was used to assign IDs depending on factors such as distance, a person's distinctive appearance, and the likelihood of their next position (current speed and direction of movement are taken into account). This is how the video looks like after the ML models and Deep Sort are applied:

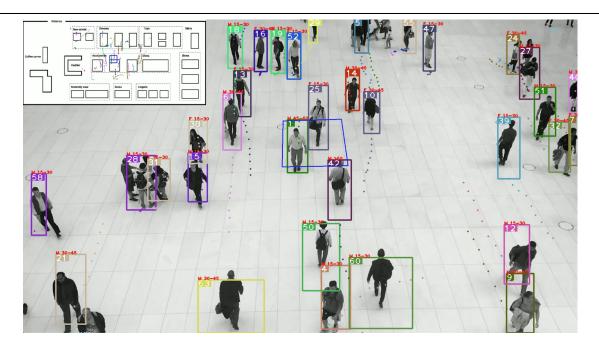
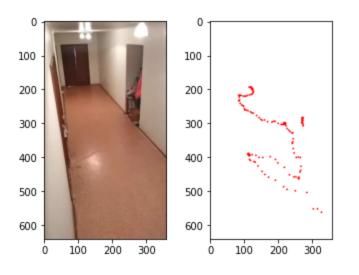


Figure 6. Video from a CCTV camera with image recognition applied.

#### **Projection of coordinates**

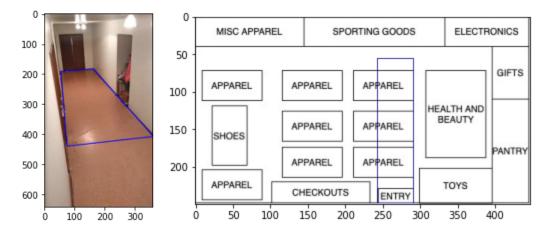
To create a heatmap for in-store behavior analytics, a projection on a collection of *x*, *y* points was chosen. It is a lot faster than image manipulation. The flow from camera to the UI is implemented as follows:

1. Collecting *x*,*y* coordinates of each person from the camera video feed and sending their recorded position as a collection to the relational database (MySQL).



*Figure 7.* Illustration of video projection to a set of *x*, *y* coordinates.





*Figure 8.* Mapping the plane from camera angle on the store layout.

3. Requesting all coordinates for a time range and projecting them on the layout.

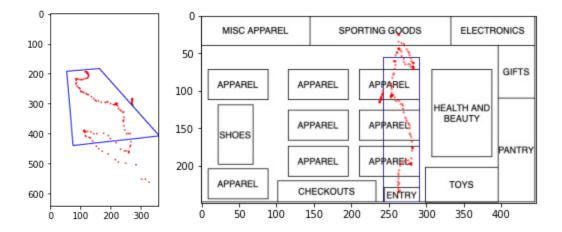
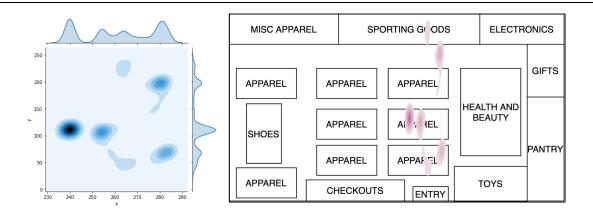
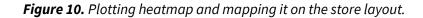


Figure 9. Translating the set of coordinates to the store layout.

4. Plotting heatmap using scatter density plot of requested coordinates, then aligning and mapping it on the store layout.





Use Case Highlight	
Using Dell EMC PowerEdge XE2420 for real-time video recognition and processing: CPU/GPU comparison.	<ul> <li>Taking 2 FPS (frames per second) as the controlled latency, sufficient for real-time tracking for the use cases A and B, here's what the average processed FPS data looks like: <ul> <li>GPU: 5.945081 (real-time)</li> <li>CPU: 1.437055 (near real-time)</li> </ul> </li> <li>Using GPU allows for a real-time performance that, paired with high local bandwidth (10GBps to 100GBps) proves to efficiently handle the ML workloads. Moreover, less infrastructural support costs, better security, and higher analysis quality due to no data compression, make edge computing on Dell EMC PowerEdge XE2420 a great plug &amp; play environment for retail AI systems.</li> </ul>

Use Case 3: Smart inventory management

Use case 3 simulates in-store sales and replenishment process with regard to adjusting the inventory and store assortment. It provides a simulated catalog, SKUs prices and their daily distribution, customer preferences, demand forecast, weekly stock level updates based on the forecast, weekly replenishment based on availability and stock level.

The tool analyzes actual and forecasted demand, visualizes it on a simple dashboard and provides alerts on differences for future corrections.

Key services and technologies

Use Case 3: Python, Apache Airflow, Elasticsearch, Pandas, Scikit-learn

Visualizing

The tool fetches data for visualization from Elasticsearch using REST API. Alerts are constructed for the fetched data on the backend.

#### Simulating data

Catalog Generation	Using elasticsearch index c_catalog, which is filled with documents:
	• id =020200001, name (type:keyword) = SKU 020200001
	• id =020200002, name (type:keyword) = SKU 020200002
	•
	• id =020200020, name (type:keyword) = SKU 020200020
	Number of SKUs is configurable (currently 20).
Price generation	Prices are generated using the following formula:
	$P_p(t) = a_p \sin(\alpha_p t + \beta_p) + a_p + v_p$ , where
	<ul> <li>t is date</li> <li>{a<sub>p</sub>}<sup>20</sup><sub>p=1</sub>, {α<sub>p</sub>}<sup>20</sup><sub>p=1</sub>, {β<sub>p</sub>}<sup>20</sup><sub>p=1</sub>, {v<sub>p</sub>}<sup>20</sup><sub>p=1</sub> are sequences generated from normal distribution (for example μ = 50, σ = 25, but parameters are configurable) for the list of SKUs ordered by id.</li> </ul>
Customer preferences generation	Preference ( $C_p(t)$ ) is coefficient from [0.1, 1] which depends on a day of week and is applied for the generated demand. Weekend has greater preferences than work-days for 70% of SKUs, and vice versa for other 30% of SKUs.
SKU availability	SKU availability is the number of SKU items on the store front (shelve):
	$A_p(T_{Sunday} + \tau) = max(A_p(T_{Sunday} + \tau - 1) - d_p(T_{Sunday} + \tau - 1), 0)$ where:
	• $\tau = 1, 2, 3, 4, 5, 6 A_p(T_{Sunday}) = \sum_{\tau=0}^{6} I_p(T_{Sunday} + \tau)$
	<ul> <li>d<sub>p</sub>(t) is actual demand (sold item by day t)</li> <li>I (t) is stock level for day t</li> </ul>
	<ul> <li>I<sub>p</sub>(t) is stock level for day t</li> <li>I<sub>p</sub>(t) = A d         <ul> <li>I<sub>p</sub>(t) = A d             <li>I<sub>p</sub>(t) + B, d             </li></li></ul> <li>I<sub>p</sub>(t) = A d             </li> </li></ul>
	configurable parameters.

Demand generation	Actual demand: $d_p(t) = min(E(P_p(t)) * C_p(t), A_p(t))$
	E(x) is an elasticity curve that provides demand dependency on price $x$ .
	logE(x) = log S - L log x, where
	<ul> <li><i>x</i> is price</li> <li><i>L</i> is elasticity coefficient</li> </ul>
	Setting $S$ and $L$ and calculating the elasticity curve.

#### **Demand forecast model**

Using formulas for simulating data, generating enough historical data for a significant amount of time (2 years at least) and continuously training a demand forecast linear model.

#### Workflow:

- 1. Preparing configuration (coefficients) for the historical dataset.
- 2. Generating historical data and storing it in the Elasticsearch indices.
- 3. Training demand forecasting model.
- 4. Generating historical forecasts and storing them in the Elasticsearch index.
- 5. Scheduling a daily job to generate data for a current day and the next day, getting a forecast for the number of days until the next Saturday (inclusively), and storing it in the Elasticsearch indices.
- 6. Using data collected from the Elasticsearch, calculating difference (RMSE) metric and showing an alert if the threshold is exceeded.

#### **Use Case Highlight**

Using Dell EMC PowerEdge XE2420 as an edge-computing endpoint for retail analytics. Performing end-to-end AI-driven analytics on Dell EMC PowerEdge XE2420 has proven to have the following benefits:

- **Focused analytics**: using local data for forecasts allows translating into actionable insights for a particular store while keeping the overall system unified across all stores.
- **Data granularity**: having volumes of data analyzed locally in real time, edge processing saves both resources and time needed to transfer it to the centralized system. More granular and timely interpretation therefore ensures precision and value of the results.
- **Autonomous stability**: all the insights are available regardless of any issues with accessing the centralized system.

In omnichannel retail solutions, a combination of centralised corporate systems and local systems would likely be used, allowing improvement in overall strategy and fine-tuning at the ultra-local level.

#### Use Case 4: Detecting anomalies on the server

Any analytics, AI-powered in particular, requires and generates huge amounts of data. As downtime or any issues with equipment affect business, this use case concentrates on detecting and predicting any abnormal behavior of the server. Using out-of-the-box iDRAC telemetry endpoints, the anomaly detection use case shows a prototype for catching suddenly occurring outliers and also revealing changes in the distribution of very noisy data. The solution is AI driven and implements a flexible approach based on normal state and behaviour patterns extraction, but does not rely on purely statistical methods.

Key services and technologies	
Use Case 4:	Apache Airflow, Elasticsearch, MySQL, HTM models, Flask, AngularJS

#### Data scheme

The data scheme consists of: metric name, time, value, metric source. We simulate data using a configurable distribution for every time unit of each metric (for example, 1 minute).

#### Anomaly detection flow

Anomaly detection tasks comprise:

- data transformation and adoption;
- HTM model invocation for real-time training and inference;
- anomaly detection decision making.

HTM Endpoint encapsulates HTM models for all metrics. Models are stored and change their state in memory and not persisted. It means that there is no automated failover to recover trained models. So automated persistence, export and import of features for models are out of scope. At the same time, we provide HTM Endpoint API to download and upload models.

Alerting DAG is a pipeline for fetching the detected abnormal use cases from ES indices, applying alerting business rules, and sending emails if rules are satisfied.

#### Components

Data organization	Using Elasticsearch 7.x single-node cluster with one shard per each index
	and no replicas. Indices are used to store:
	<ul> <li>ad_metric — generated telemetry data metrics.</li> </ul>
	<ul> <li>ad_anomaly — detected anomalies.</li> </ul>

	<ul> <li>ad_scores — anomaly detection anomaly scores and likelihoods.</li> </ul>
Anomaly detection configuration	<ul> <li>Storing the configuration in MySQL, CE, 8.x single node database. This includes:</li> <li>data ingestion configuration</li> <li>anomaly detection configuration</li> <li>alerting configuration</li> </ul>
Simulating data	Using Airflow DAG to generate simulated metrics and abnormal use cases.
Visualizing	<ul> <li>All configuration is visualized on the Anomaly Detection Dashboard, which includes:</li> <li>Configuration of HTM Models and thresholds (simple SDR configuration: min/max; Likelihood threshold, metric).</li> <li>Switching the HTM model training on/off.</li> <li>Alerting Configuration via a business rule using a number of abnormal points and number of unique abnormal metrics; mailing list.</li> <li>Metrics time series visualization and anomalies highlights.</li> <li>Anomalies statistics.</li> </ul>

Use Case Highlight	
Using iDRAC9 to detect anomalies on Dell EMC PowerEdge XE2420.	Using telemetry endpoints from integrated Dell Remote Access Controller (iDRAC) is sufficient for training HTM models and building a solution that provides unsupervised learning and anomaly detection for server metrics.

## Reference architecture model and configuration

#### **Server Configuration Used**

The Dell Retail Edge Analytics system reference architecture is developed on the Dell EMC PowerEdge XE2420 server with the following specifications:

Processor	2x Intel Xeon Gold 5220 CPU
RAM	128 GB
CPU	2x Intel Xeon Gold 5220 CPU @ 2.20GHz

GPU	NVIDIA Corporation TU104GL [Tesla T4] (rev a1)
Storage	512 GB
Network	Broadcom Gigabit Ethernet BCM5720
OS	Linux Ubuntu 20.04 Server Edition

#### **Deployment process**

- 1. Install the PowerEdge server and OS with link to Server's user guide.
- 2. Install a local Jenkins instance on the server itself or a Jenkins agent that connects to a remote Jenkins instance.
- 3. Configure a working Kubernetes cluster: single-node K3s set up on the XE2420 server.
- 4. Cloning the repo with the reference architecture.
- 5. Deploy shared service charts to run additional components (Airflow, MySQL, Elasticsearch, Grafana) on top of Kubernetes using the scripts from the repo.
- 6. Generate historical data used for training the AI models using a generation pipeline built with Airflow DAGs, and then put the historical data into the corresponding data source (MySQL, Elasticsearch).
- 7. Build service docker container images for the use cases (7 in total).
- 8. Configure Ingress route/configuration for frontend-backend communication.

#### **Reference Architecture**

The reference architecture is designed to simulate processing all of the data and train ML models directly on the Dell EMC PowerEdge XE2420 server. As decreasing data latency and improving the speed of converting it into actionable insights is one of the key business values of edge computing, the system simulates the capability of Dell EMC PowerEdge XE2420 to serve as an end-to-end environment for handling significant AI workload.

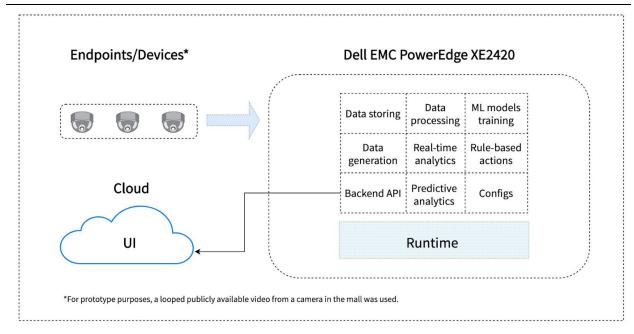
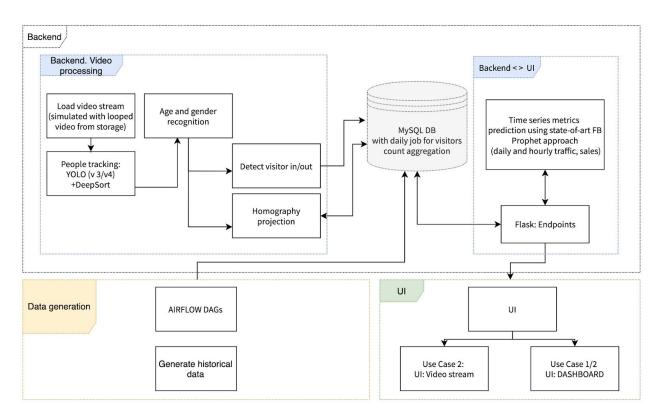


Figure 11. High-level architecture of Dell Retail Edge reference architecture.

Benefits of using Dell EMC PowerEdge XE2420 as the edge server for retail

Typical latency demonstrated:: • Edge server: < 1 ms • Public cloud: <50-55 ms	<ul> <li>In addition to drastically reducing data latency, using edge servers also introduces plenty of potential benefits for retail business.</li> <li>The most prominent ones include: <ul> <li>bandwidth availability and use;</li> <li>network connectivity (no disruptions) and security;</li> <li>autonomy and data privacy.</li> <li>With up to two Intel Xeon processors available, the XE2420 delivers the performance levels required for demanding edge environments.</li> </ul> </li> <li>Beyond that, edge servers provide better precision of results by processing and prioritizing data as it comes.</li> </ul>
---	--

#### Use case architecture



*Figure 12.* High-level architecture for eCommerce-style analytics for Brick & Mortar stores and advanced in-store behavior analytics use cases.

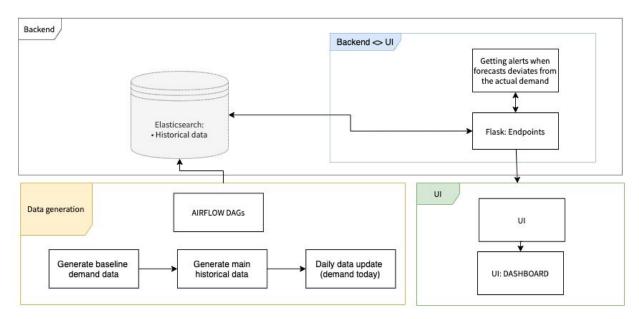
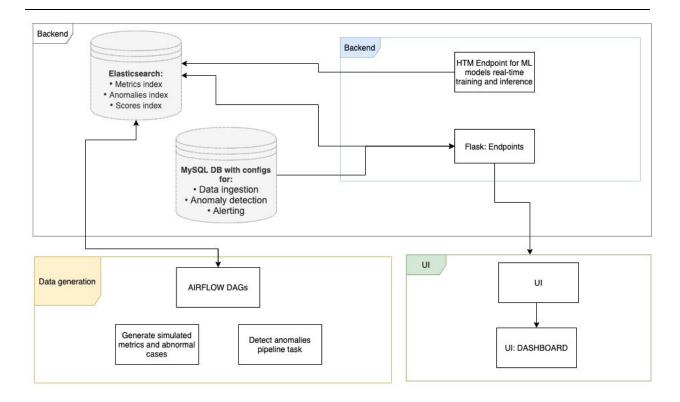


Figure 13. High-level architecture for inventory decision support tool use case.



#### Figure 14. High-level architecture for anomaly detection in server metrics use case.

For ease of access, the UI is stored on Google Cloud and is shared between all the use cases. Each use case is available in a separate tab. The communication between frontend and backend is run using Flask Rest API with GET requests. Each request is used to generate a separate dashboard widget. All requests follow / [case\_name] / [request] syntax.

#### Conclusion

The scope of the data used in the reference architecture validates that Dell EMC PowerEdge XE2420 handles large amounts of operational AI-related workload.

Designed for Edge specific use cases, PowerEdge XE2420 Server provides you optimized computing power that includes CPUs and GPUs. Combined with high performance direct attached storage, sitting locally at the Edge, you can leverage a high bandwidth network to capture high quality data and analyze it locally to gain insight into your business. This enables your Global business with local insights.

With regards to scalability and performance capabilities of Dell EMC PowerEdge XE2420 in object detection, it is expected that if the required extensions are implemented to support multiple cameras and one GPU-consuming service will process videos from all streams in batches, performance should still be in real-time.

#### About 2<sup>nd</sup> Generation Intel Xeon Scalable processors

The latest from Intel, the 2nd Generation Intel Xeon Scalable processor platform features a wide range of processors to support the workloads you run, including Bronze, Silver, Gold, and Platinum. According to Intel, the 2nd Generation Intel Xeon Scalable platform can handle a variety of workloads, including enterprise, cloud, HPC, storage, and communications.<sup>1</sup>

<sup>1</sup>Intel, "2nd Gen Intel Xeon Scalable Processors Brief," accessed October 8, 2019, <u>https://www.intel.com/content/www/us/en/products/docs/processors/xeon/2nd-gen-xeon-scalable-processors-brief.html</u>.

#### About the Dell EMC PowerEdge XE2420

The Dell EMC PowerEdge XE2420 is a dense 2U, two-socket server. It features support for up to 16 DDR4 DIMMs and offers optional integrated filtration for harsh environments. To learn more about the Dell EMC PowerEdge XE2420, visit <u>https://www.dellemc.com/en-us/collaterals/unauth/data-sheets/products/servers/poweredge-xe2420-specshe</u> et.pdf

#### **About Grid Dynamics**

Grid Dynamics is a leader in driving enterprise-level digital transformation services for Fortune 1000 corporations. The company helps organizations become more agile and create innovative digital products and experiences using deep expertise in emerging technologies such as artificial intelligence, data science, cloud computing, Big Data and DevOps, top global engineering talent, lean software development practices, and a high-performance product culture.

To learn more about Grid Dynamics visit https://griddynamics.com