

Lab Insight
**Dell Delivers
CPU-Based AI PoC for
Retail – Executive
Summary**

AUTHOR **Mitch Lewis**
Research Analyst | The Futurum Group

APRIL 2024



AI for the Retail Market

The rapid development of AI technology presents a significant opportunity for data driven retail organizations to enhance customer experiences and optimize store operations with new AI powered applications. The possible applications of AI within retail environments are broad, and may include product recommendation systems, virtual assistants, fraud and theft detection systems, AI-optimized pricing systems, or inventory management.

By capturing and analyzing data with AI-driven insights and predictions, retail organizations can rapidly adapt to shifts in customer demand, purchasing behavior, logistical challenges, and market trends rather than taking a more traditional, reactionary approach. While adoption of AI may be beneficial to retail organizations, unfamiliarity with AI technology and the hardware and software components needed for deployment and optimization acts as a barrier to adoption.

The following solution demonstrates a PoC for an AI-powered retail inventory management system that can be quickly deployed and further expanded upon by retail organizations using commonly available CPU-based Dell™ hardware.

Solution Overview

To demonstrate a real-world example solution of an AI application that could be deployed in a typical retail environment, Scalers AI™, in partnership with Dell, Broadcom™, and The Futurum Group, implemented a PoC solution for an AI powered retail inventory management system. The solution is designed to capture data from store cameras and use an object-detection AI model to monitor and manage product stock levels. The solution is capable of detecting products on store shelves, keeping track of inventory, and raising alerts for low or out of stock items.

The retail management solution addresses a common challenge in retail environments of inventory distortion. Without accurate and timely inventory management, retail organizations can be challenged with stock levels that are either too low or too high. By utilizing computer vision and object detection AI models to monitor and track inventory, retailers can achieve real-time insights into their stock to balance their inventory more efficiently.

The following provides a high-level overview of the PoC solution:

- Video data of store shelves captured by in-store cameras is streamed to an AI video pipeline built for retail product detection.
- Inferencing is achieved with an SSD_Mobile_Net_v2 image detection model that has been fine tuned to detect specific retail products.
- A separate visualization service provides live visualization of the store shelves and detected items. It additionally displays item counts, and raises alerts for low inventory.
- Image data can additionally be sent to NAS storage for long term retention.

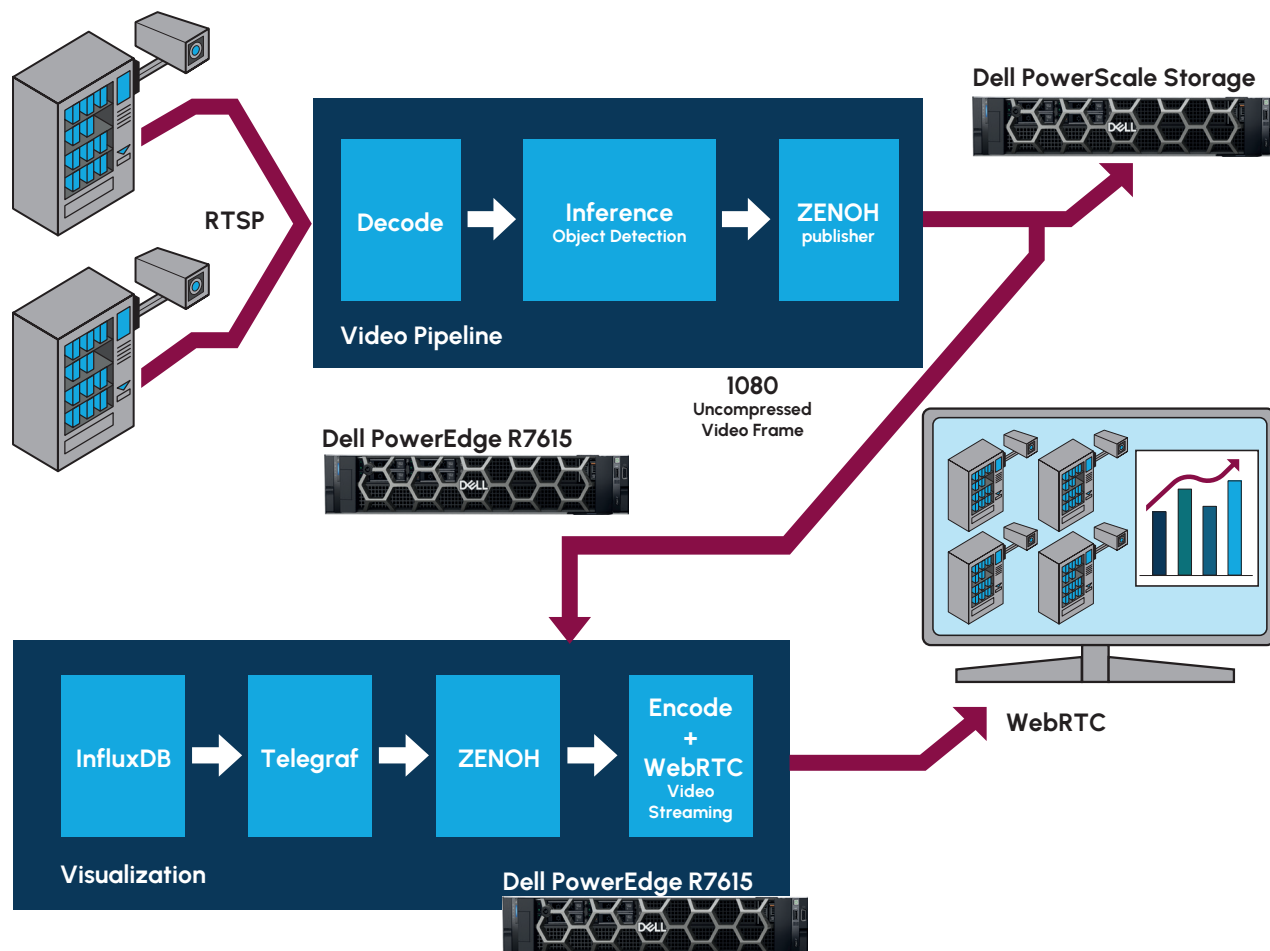


Figure 1: Retail Inventory Management AI Pipeline (Source: Scalera AI)

Key Highlights for AI Practitioners

- Dell PowerEdge R7615 Servers
 - AMD EPYC 9354P 32-Core Processors
 - 768 GB Memory
 - 1 TB Storage
 - Broadcom BCM57508 NetXtreme-E 200G Ethernet Controller
- Dell PowerSwitch Z9664
- Dell PowerScale Scale-Out NAS Storage
 - Optional for long term retention



Solution Highlights

The retail inventory management PoC solution provides valuable insight for both AI practitioners and IT operations professionals working in retail environments. Notable highlights include:

- The solution was achieved using off the rack, CPU-based Dell servers. No GPUs or specialized accelerators were required.
- Both model training and deployment of the pipeline were achieved with the same 32 core Dell PowerEdge Servers.
- The solution was optimized by utilizing AMD's ZenDNN library with node pinning. Performance testing demonstrated 1.5x throughput increase compared to the default configuration.
- The PoC solution is designed to accommodate additional scalability with more video streams, products, and application logic. Separation of the AI video pipeline and visualization service, connected by high bandwidth Broadcom Ethernet, provides flexibility for additional future scaling.
- Customization of the base SSD_Mobile_Net_V2 model was achieved with transfer learning to provide retail specific image detection.

Conclusion

The retail market presents a clear opportunity for AI-powered solutions to make an immediate impact. Limited familiarity with AI and expectations of specialized hardware requirements, however, may delay AI adoption for some organizations. The retail inventory management PoC described in this paper demonstrates the ability to quickly deploy a valuable, real-world AI solution for retail environments using off the rack, CPU-based hardware. By leveraging reference solutions such as this, retail organizations can quickly deploy AI applications to optimize operations, enhance customer experiences, and save costs.

Important Information About this Report

CONTRIBUTORS

Mitch Lewis

Research Analyst | The Futurum Group

PUBLISHER

Daniel Newman

CEO | The Futurum Group

INQUIRIES

Contact us if you would like to discuss this report and The Futurum Group will respond promptly.

CITATIONS

This paper can be cited by accredited press and analysts, but must be cited in-context, displaying author's name, author's title, and "The Futurum Group." Non-press and non-analysts must receive prior written permission by The Futurum Group for any citations.

LICENSING

This document, including any supporting materials, is owned by The Futurum Group. This publication may not be reproduced, distributed, or shared in any form without the prior written permission of The Futurum Group.

DISCLOSURES

The Futurum Group provides research, analysis, advising, and consulting to many high-tech companies, including those mentioned in this paper. No employees at the firm hold any equity positions with any companies cited in this document.



ABOUT THE FUTURUM GROUP

[The Futurum Group](#) is an independent research, analysis, and advisory firm, focused on digital innovation and market-disrupting technologies and trends. Every day our analysts, researchers, and advisors help business leaders from around the world anticipate tectonic shifts in their industries and leverage disruptive innovation to either gain or maintain a competitive advantage in their markets.



CONTACT INFORMATION

The Futurum Group LLC | futurumgroup.com | (833) 722-5337