

Zoho is optimizing and accelerating video AI workloads

Video AI Assistant by Zoho, powered by Intel’s hardware and software platforms, accelerates AI workloads resulting in improved performance and efficiency, reduced latency, energy efficiency, and lower TCO.

Authors

Vishnu Madhu
AI Software Solutions Engineer, Intel

Ritik Agrawal
Industry Technical Specialist, Intel

Raja Gopal Hari Vijay
Member - Leadership Staff, Zoho

Praveen Vijayakumar
Member - Technical Staff, Zoho

Bhargav K
Member - Technical Staff, Zoho

Sreekar G
Member - Technical Staff, Zoho



Executive Summary

Zoho, a leading technology company and India’s premier software-as-a-service provider, embarked on a strategic initiative to harness the transformative potential of artificial intelligence (AI). Recognizing the importance of AI in driving innovation and efficiency, Zoho aimed to optimize their AI workloads to meet optimal price-performance Total Cost of Ownership (TCO) requirements.

In line with this objective, Zoho collaborated closely with Intel to benchmark and optimize key AI workloads which include Closed-circuit Television (CCTV) surveillance videos analytics and Optical Character Recognition (OCR).

Through iterative benchmarking and optimization processes, Zoho and Intel fine-tuned the performance of these AI workloads, ensuring optimal resource utilization. Leveraging Intel’s expertise in hardware and software optimization, Zoho achieved tangible enhancements in both performance and cost-effectiveness across these critical AI applications.

Table of Contents

Executive Summary.....	1
The Need of Seamless Data Management for AI Workloads.....	2
Overcoming Challenges in Accelerating AI Workloads.....	2
Optimizing Video Analytics and Text Digitization with.....	2
Delivering Benefits of Accelerated and Optimized AI Workloads.....	3
Maximizing Intel’s AI Accelerators for Accelerated Outcomes.....	3
Gaining the Intel Edge.....	4
Enabling Real-world Impact.....	4
Conclusion.....	5

“ At Zoho we collaborate with leading technology industry vendors in improving the TCO of our Infrastructure and solutions to ensure we offer the best value to our customers.

In the same line, we are working with Intel in optimizing and identifying the best case TCO for our AI service offering needs. We have an ongoing relationship with Intel across platform generations and looking forward to bring best Performance/Dollar solutions for the AI and Analytics domain as well. ”

Shailesh Kumar Davey
Co-Founder, Vice President,
General Management-Engineering,
Zoho

The Need of Seamless Data Management for AI Workloads

Video analytics and OCR are two key areas where AI is making significant strides. These processes often involve handling enormous amounts of data and complex computations.

Video analytics using CCTV surveillance cameras is a multifaceted process that involves the real-time or near real-time processing of vast quantities of video data. This demanding task requires the implementation of sophisticated algorithms capable of handling various tasks including enhancing recording quality, detecting objects, and tracking individuals. Similarly, Tesseract OCR, an OCR engine, plays a key role in the digitization of text from scanned images, video frames, or documents while supporting multiple global languages.

In applications such as surveillance, document digitization, and text analysis where rapid decision-making and real-time processing are critical, hardware acceleration as part of conventional compute engines offers significant benefits.

By using these built-in accelerators within the CPU cores computational tasks, organizations can achieve optimal TCO for faster processing speeds, reduced latency, and improved scalability. The same computing platforms can also be repurposed for general purpose workloads, ensuring optimal use of the platform infrastructure.

Overcoming Challenges in Accelerating AI Workloads

AI provides a competitive advantage for businesses, however accelerating AI workloads presents a tough challenge, particularly as organizations strive to reduce latency, improve energy efficiency, and optimize TCO. At the outset, processing large volumes of data is a significant challenge, particularly when dealing with diverse document sizes and live video streams from a multitude of CCTV cameras, spanning thousands of cameras and a range of resolutions from HD to 4K (Refer Figure 1). Managing this data influx requires robust infrastructure and efficient data processing pipelines.

Camera Resolution	Camera Count
1280 * 720 (Approximately HD or 720p)	381
1280 * 960	11
1920 * 1080 (Full HD or 1080p)	1315
2560 * 1440 (2K or Quad HD)	23
3072 * 1728	1
3820 * 2160 (4K or Ultra HD)	7
Total	1738

Figure 1: Details of cameras used by Zoho on a weekly basis for video analytics

Furthermore, the complexity of AI and image processing algorithms compounds the challenge. Converting video streams into individual frames, resizing images, removing noise, and enhancing image quality are just a few of the tasks necessary for AI inferencing. These operations demand substantial computational resources and sophisticated algorithms to ensure accurate and timely results.

Another key challenge to consider is the TCO, which is bringing AI solutions to cost-conscious end customer markets. Striking a delicate balance between performance and affordability is key for successful deployment and consumption of AI services, and meeting best-case performance-to-cost thresholds is essential to ensure widespread adoption and customer satisfaction.

Additionally, integrating AI solutions with existing systems poses another obstacle. Streamlining data flow, enhancing interoperability, and improving overall system performance are critical considerations to ensure seamless integration and maximize the value derived from AI implementations. In addition to the aforementioned challenges, it's crucial to acknowledge and address other significant concerns pertaining to cost, scalability, and energy efficiency.

Optimizing Video Analytics and Text Digitization

Zoho's Video AI solution, powered by Intel's hardware and software platforms, serves as a key tool in enhancing physical security and safety in the premises and employees. This underlying platform includes Intel® Xeon® Scalable processors and Intel® Distribution of OpenVINO™ toolkit. This solution enhances performance, provides quicker insights, and enables more streamlined processes in applications for surveillance, document digitization, and text analysis.

Video analytics is the key capability powering the Video AI Assistant. The Camera Image Quality Analyzer is a key

module of the tool that helps identify cameras with sub-optimal recording quality due to environmental factors such as dust, fog, or spider webs. Post this, the system tickets the issue and reports for tracking and resolution. Identifying and rectifying these issues ensures the capture of high-quality video footage, thus enhancing the accuracy of subsequent incident analysis. Moreover, the Video AI Assistant facilitates effective resource management through its people counting functionality - enhancing operational efficiency and contributes to a more organized workplace environment. In large spaces where sensor sensitivity may be reduced, the Video AI Assistant can count people using the camera feed and adjust the air conditioning accordingly. This temperature conditioning application enhances efficiency and conserves energy, demonstrating the diverse capabilities of AI in optimizing operational efficiency and improving sustainability.

On the text digitization front, Tesseract OCR serves as a fundamental component of Zoho's AI Assistant, driving the digitization of text to identify key topics and content within documents. This enables the AI assistant to generate relevant search terms, facilitating streamlined document management and information retrieval processes. Each day, a significant volume of documents undergo OCR processing, numbering in the thousands. However, due to the sheer magnitude of this workload, there are instances where completion of processing extends beyond the standard timeframe, occasionally necessitating an additional day for the processing of remaining documents. By using the right Intel® Xeon® features and associated software optimizations, the challenges are mitigated.

Delivering Benefits of Accelerated and Optimized AI Workloads

Zoho's Video AI Assistant can significantly reduce processing time, simplify deployment, and accelerate workloads through optimal utilization of Intel's hardware and software platforms. Some of the key benefits of the Intel-powered Video AI Assistant solution include:

- **Timely Decision-Making for Real-time Resource Management:** AI Assistant provides actionable insights to optimize resource allocation, including dining hall capacity management and equipment utilization, enabling efficient operations and cost savings.
- **Solution Portability:** Whether deployed on-premises or in the cloud, the solution maintains consistency in data and deployment workflows, ensuring seamless accessibility and continuity across environments.

- **Simplified Deployment:** Integration complexities are minimized through Intel-enabled pre-configured connectors, package libraries, and deployment templates. This streamlined the implementation process and accelerated time to value.

- **Integration with Zoho Products and IoT:** By seamlessly integrating with Zoho products and IoT devices, Video AI Assistant unlocks new possibilities for data-driven insights and automation, enhancing overall efficiency and effectiveness.

- **Unified User Experience:** AI-driven insights are consolidated within familiar Zoho interfaces, eliminating the need for context switching and streamlining decision-making processes for enhanced productivity and user satisfaction.

Maximizing Intel's AI Accelerators for Accelerated Outcomes

Zoho recognized the necessity for a robust platform capable of meeting the high-performance demands of its Video AI Assistant solution. By collaborating with Intel, Zoho optimized the Video AI Assistant platform by taking advantage of Intel's hardware and software architecture including:

- **Intel® Xeon® Scalable Processors:** AI processing can be a very expensive and power-intensive task, especially in the case of 24/7 computer vision monitoring and detection. Intel® processors help make AI much more affordable and accessible. Video AI Assistant runs resource-intensive workloads exceptionally well on Intel® Xeon® processors, leveraging built-in AI acceleration features.

- **Intel® Advanced Matrix Extensions (Intel® AMX):** To optimize and accelerate AI pipelines and workloads, Zoho's Video AI Assistant maximizes 4th Gen Intel® Xeon® processors with Intel Advanced Matrix Extensions (Intel AMX), a built-in AI accelerator. Intel AMX is leveraged by Zoho's Video AI Assistant for the Deep Neural Networks.

- **Intel® Distribution of OpenVINO™ Toolkit:** Intel's Open Visual Inference and Neural Network Optimization Toolkit enabled Zoho to build and optimize the AI models for the Video AI Assistant plug-and-play library, ensuring best AI performance on Intel's hardware. Zoho's Video AI Assistant solution harnesses the power of OpenVINO™ runtime in conjunction with the Neural Network Compression Framework (NNCF) compression technique. NNCF offers a sophisticated methodology for compressing neural network models, reducing their computational complexity and

memory footprint without sacrificing accuracy. By applying NNCF compression techniques, Zoho is able to optimize the efficiency of its AI models, improving inference speed and resource utilization while maintaining high levels of accuracy.

- **Intel® Distribution for Python:** Intel® Distribution for Python, a performance-optimized distribution of Python that takes full advantage of Intel’s hardware. The Video AI Assistant solution utilizes SciKit-Learn Library in conjunction with Intel® Distribution for Python for enhanced performance in model training and prediction tasks.

- **IQA-Pytorch Extension with BRISQUE (NR):** For image quality assessment, Video AI Assistant maximizes IQA-Pytorch Extension with BRISQUE (NR). By assessing image quality without the need for reference images and effectively identifying common distortions, it can enhance the performance of applications reliant on high-quality image data.

Gaining the Edge

Intel’s hardware and software laid the foundation of providing a myriad of benefits for Zoho’s Video AI Assistant that include:

- **Architecture-specific Runtime and Optimizations – Performance and Power Efficiency:** Intel offers tailored runtime support and optimizations designed to maximize the performance and efficiency of AI workloads on its hardware architecture. This specialized support ensures that Zoho’s Video AI Assistant can leverage the full potential of Intel® processors, resulting in enhanced speed, accuracy, and responsiveness in AI-driven tasks. This enhancement in speed results in reduction in time of execution significantly, thereby reducing the overall energy consumption per inference.

- **Integration with AI Open-Source Frameworks – Faster Go-to Market:** Intel’s libraries and optimizations are seamlessly integrated into numerous AI open-source frameworks, providing developers with access to a rich ecosystem of tools and resources. By leveraging these resources, Zoho was able to accelerate the development and deployment of its Video AI Assistant solution, tapping into a wealth of pre-existing functionality and expertise within the AI community.

- **Lower Total Cost of Ownership (TCO):** Intel’s hardware architecture offers a competitive TCO by optimizing

performance and efficiency. Intel processors deliver superior performance per watt, resulting in reduced energy consumption and operational costs over the lifetime of the hardware.

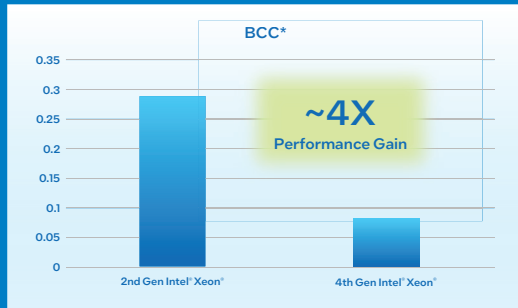
Enabling Real-world Impact

Utilizing the advantages provided by the 4th Gen Intel® Xeon® processors, Zoho has seen significant performance improvements for specific AI workloads. Below is an overview of the impact seen from the benchmarking process undertaken by Zoho.



Basic Crowd Counting (BCC) Benchmark

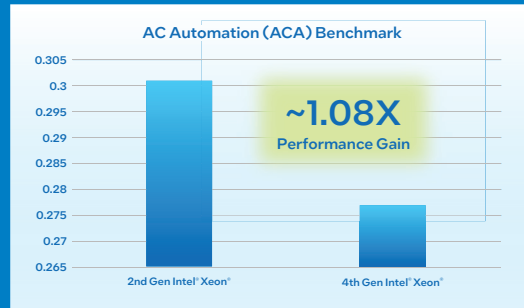
**Lower is better*



Benchmark Details: Calculation of crowds covering over 4 lunch halls. Around 9 cameras were used on a daily basis and integrated with Zoho Cliq.

AC Automation (ACA) Benchmark

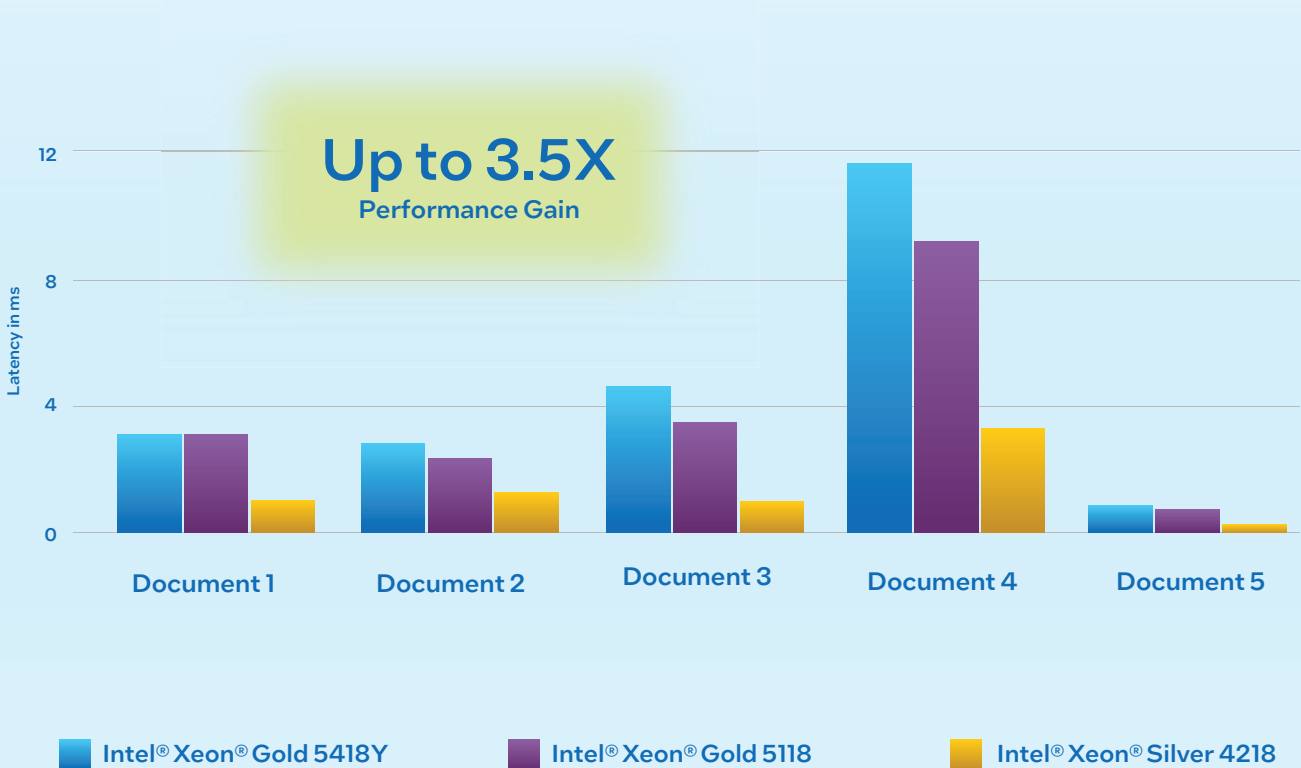
**Lower is better*



Benchmark Details: Monitoring AHUs of AC based on OpenVINO™ Model detection. Cameras were Integrated as API service and integrated with Zoho IoT.

TESSERACT OCR WORKLOAD*

**Lower is better*



Conclusion

Zoho's Video AI Assistant harnesses cutting-edge hardware accelerators like Intel AMX in 4th Gen Intel® Xeon® processors to elevate performance and efficiency, effectively tackling the complexities of processing vast datasets from CCTV cameras and OCR applications. Its seamless integration with Zoho products and IoT devices enables actionable insights for real-time resource management and decision-making, all while preserving solution portability and delivering a cohesive user experience within familiar Zoho interfaces.

Moreover, the product's integration with Intel's hardware and software platforms, including Intel® Xeon® processors and the OpenVINO™ runtime, further optimizes performance and drives down total cost of ownership. This comprehensive approach provides an ideal solution for accelerating AI workloads and overcoming challenges across diverse applications.

Spotlight on Zoho

With over 100 million users worldwide and 55+ apps in nearly every major business category, Zoho Corporation is one of the world's most prolific technology companies. Headquartered in Chennai, India, Zoho is privately held and profitable with more than 15,000 employees.

For more information, please visit: www.zoho.com.



*Internal benchmarking results from tests conducted by Zoho.

Performance varies by use, configuration and other factors. Learn more at <https://intel.com/benchmarks>

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates.

No product or component can be absolutely secure.

Your costs and results may vary.

Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

Intel technologies may require enabled hardware, software or service activation.

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries.

Other names and brands may be claimed as the property of others.