



Klepsydra AI

Efficient Software for Edge Computing

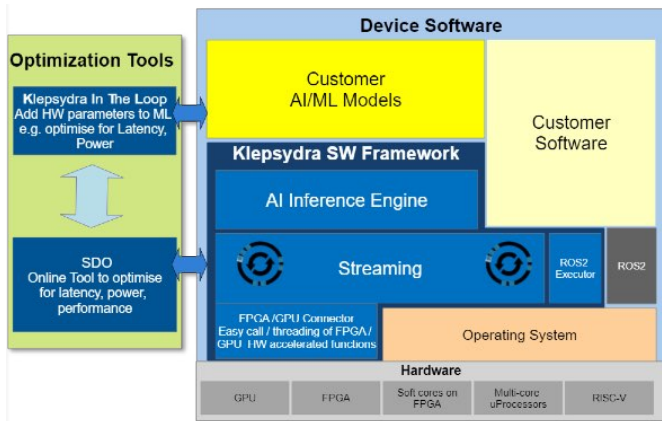


Product catalog and Technical specs
Version: 20.2.0 September 2025

Core Features

Overview

Klepsydra AI is a high-performance deep neural network engine for edge computers. It enables customers to deploy existing or newly trained models just like with standard edge AI solutions —**only with a major advantage: up to 4× higher data throughput.**



Compatibility Matrix

- **Among the networks already validated:** AlexNet, MobileNet V1 and V2, YOLOX, YOLOv4-v8 (S and N variants), ResNet50, VGG19, UNet, and LSTM.
- **Hardware and OS Compatibility:** x86, ARM Cortex-A, RISC-V, and SPARC processors, and Linux, RTEMS6, VxWorks, and DDC-I operating systems.
- **GPU Support:** NVIDIA GPU with CUDA kernels, CUDA Graphs, and TensorRT workloads.

A detailed support matrix is available upon request

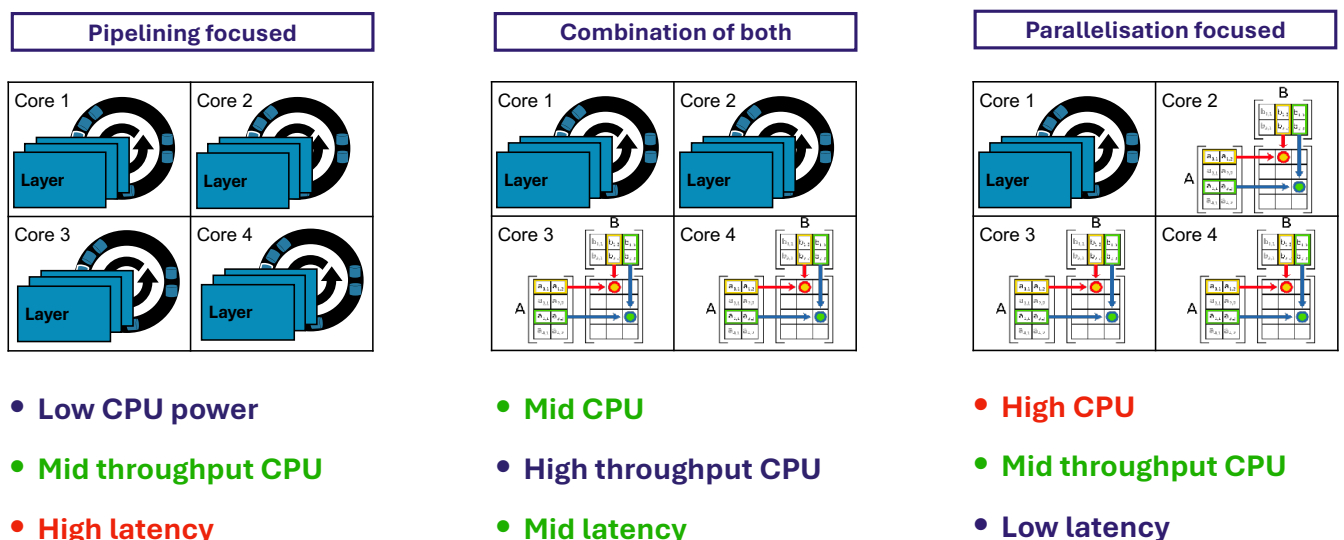
Core Technology

- **Performance Improvement:** 2× to 8× faster data processing and 75% power reduction.
- **Key Techniques:** Lock-free ring buffers, FPU vectorisation, and dedicated ring buffers per thread.
- **Optimisation:** Smart thread configuration for low latency, maximum throughput, or minimal CPU usage.

SDO (Streaming Distribution Optimiser)

- **Tool Functionality:** Optimises AI performance on target processors.
- **Optimisation Features:** Allows customisation for latency, throughput, or resource constraints.
- **Optimisation Process:** Cloud-based software collects data, runs an optimiser, and deploys the best configuration.

Klepsydra Threading model operational modes

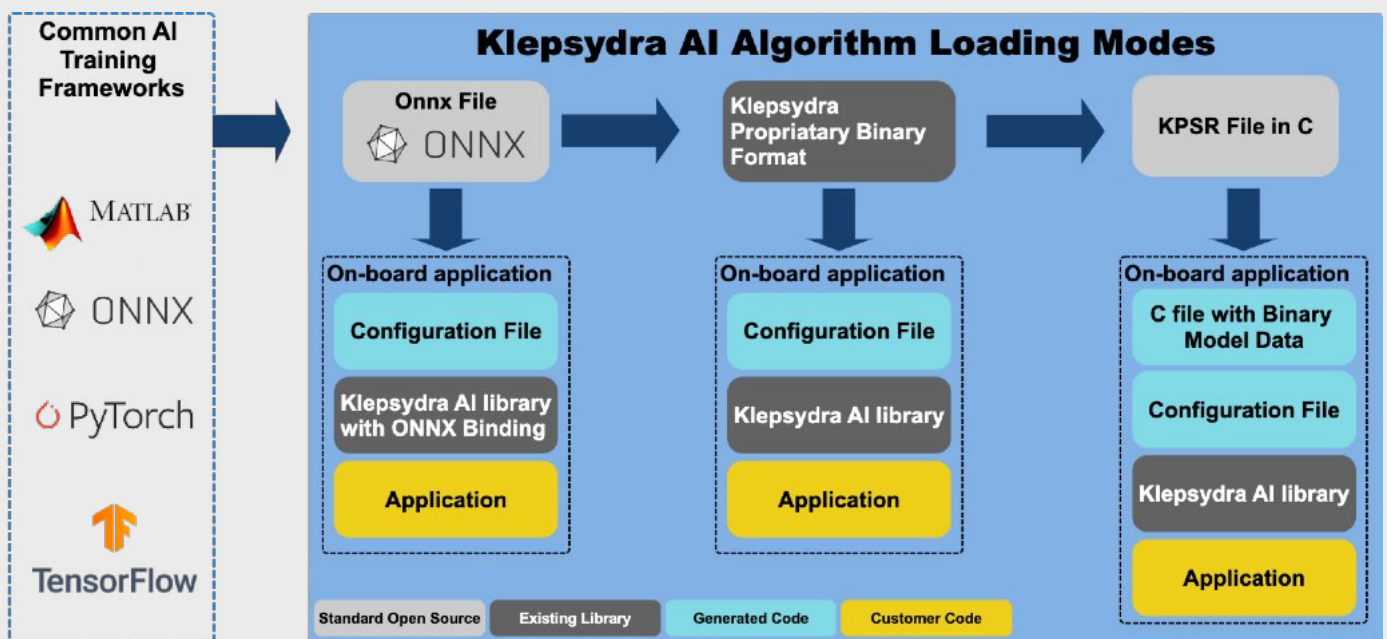
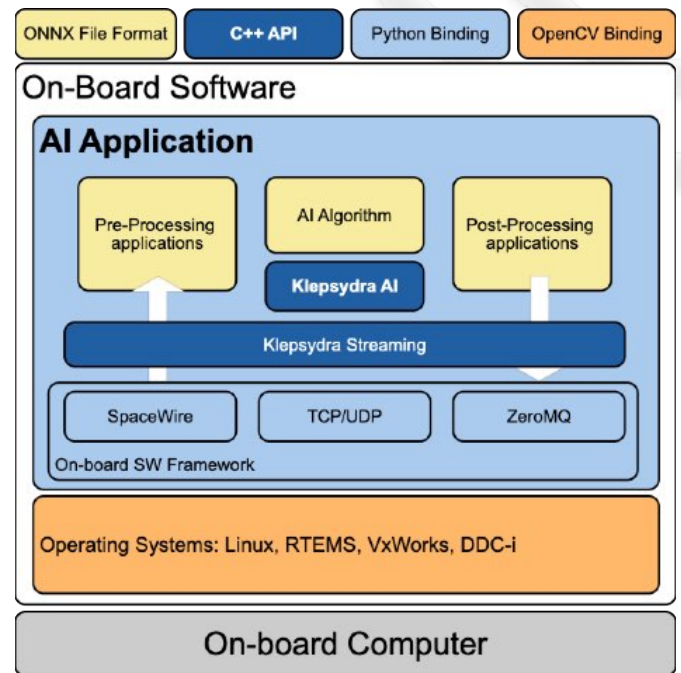


Integration with Leading AI tools

Programming Languages

- **Programming Language:** The native Klepsydra AI API is implemented in C++.
- **Bindings:** Python and OpenCV bindings are provided as open-source components.
- **Examples:** Each release includes API and binding examples to facilitate integration.

The native Klepsydra AI API is implemented in C++, with additional Python and OpenCV bindings built on top of the core API. These bindings are provided as open-source components, allowing developers to compile them with their preferred versions of Python and OpenCV.



Machine Learning Frameworks

- **Supported DNN Frameworks:** TensorFlow, PyTorch, and ONNX.
- **Model Requirements:** Export to ONNX format (Opset 12) using a channel-last approach.
- **Deployment Guide:** Klepsydra AI ONNX User Guide provides instructions for model preparation and deployment.

Machine Learning File Formats

- **Supported Formats:** Klepsydra AI supports ONNX and its proprietary binary format KPSR.
- **KPSR Format Advantages:** KPSR offers faster loading times and compatibility with platforms with limited OS support.
- **Embedding Support:** Klepsydra AI allows embedding the binary format directly into C source files for enhanced speed and security in real-time and cybersecurity-sensitive applications.

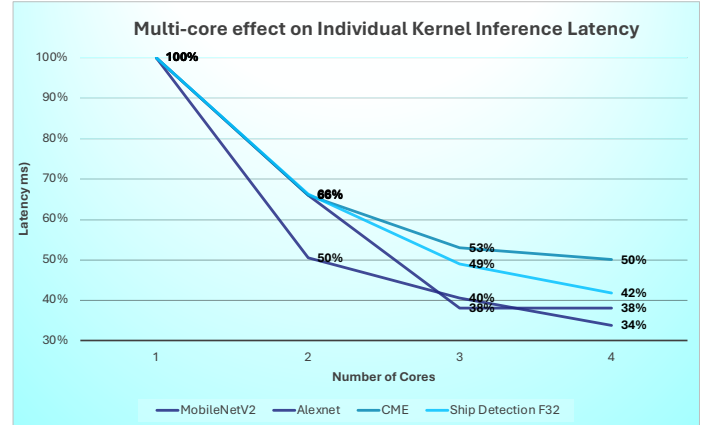
Safety and Security Features

Safety Features

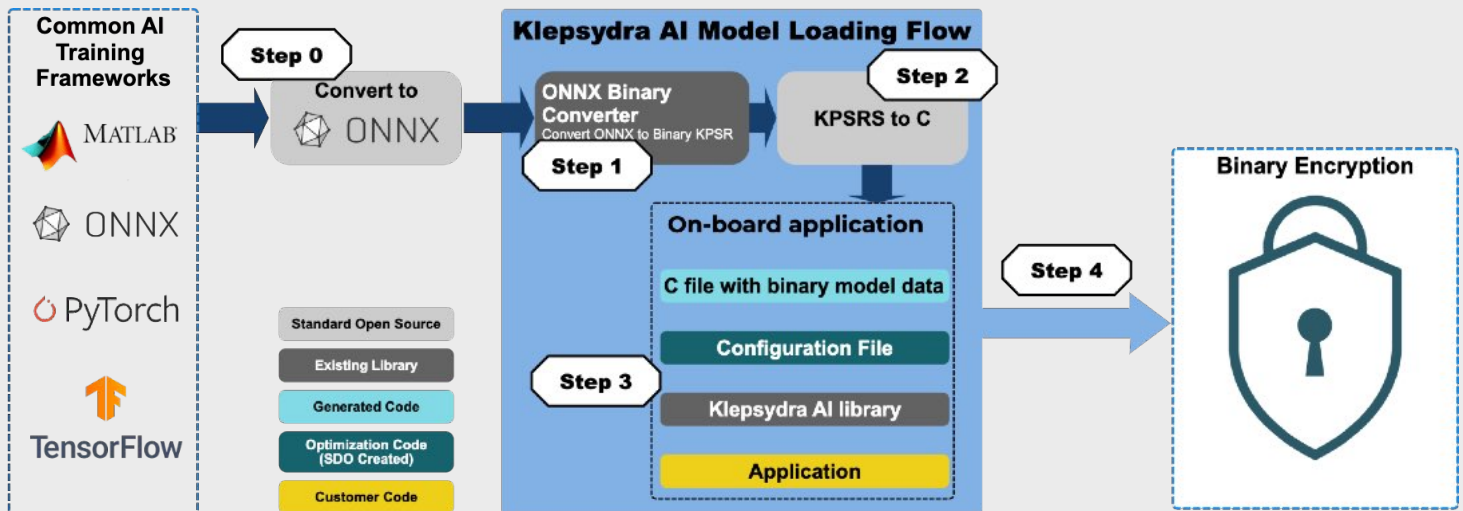
- **Standards Compliance:** Conforms to ECSS software standards E-40-ST and Q-80-ST at criticality level D, with plans for levels C and B by 2026 and 2027 respectively.
- **Operating System Support:** Supports safety-critical operating systems including VxWorks, RTEMS6 SMP, and DDC-I Deos on various architectures.
- **Processor and Hardware Support:** Compatible with space-grade processors like Gaisler, Microchip PolarFire, and ARM-based processors, along with hardware accelerators.

Multi-core Support

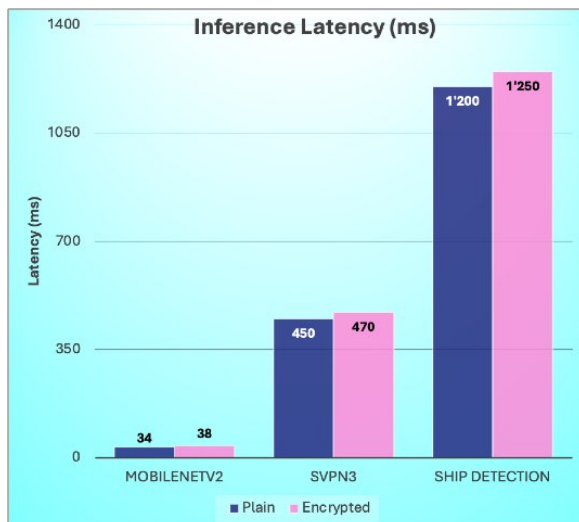
On RTEMS6 SMP and VxWorks, Klepsydra AI enables efficient multi-core execution, with performance validated by the results shown below.



Multi-core performance



Cybersecurity Protection Deployment Flow



Plain vs Secured Performance Comparison

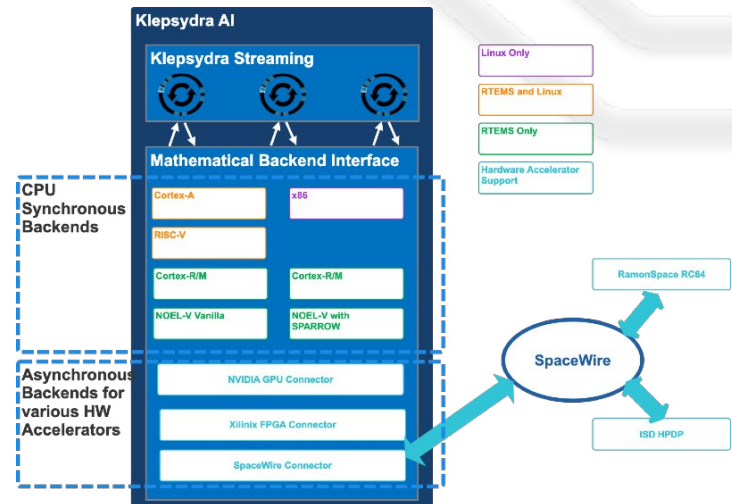
Highest Level of Cyber Protection

- **Cybersecurity Protection:** Klepsydra AI can be equipped with algorithm encryption and run-time protection to prevent reverse engineering and unauthorised modifications.
- **Performance Impact:** Negligible impact on performance, as demonstrated by benchmarks on a Raspberry Pi 4.
- **Encryption Procedure:** See Figure 6

Extensibility Framework

Overview

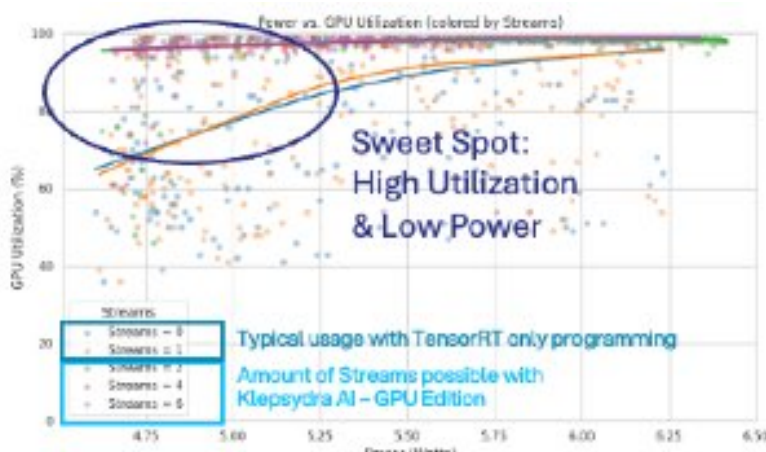
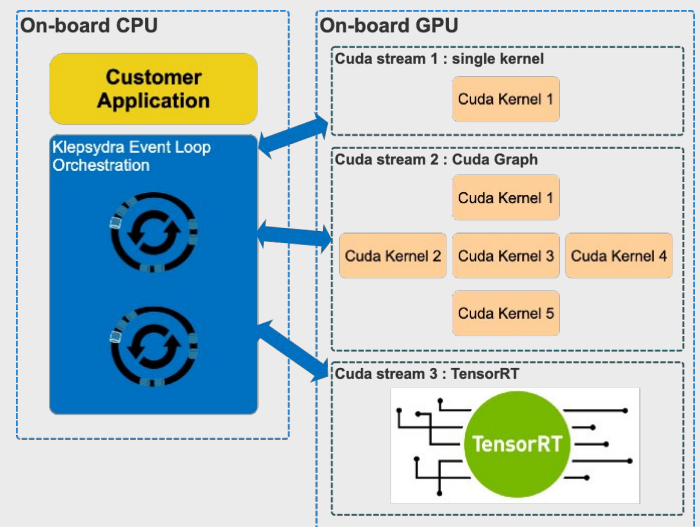
- **Extensibility:** Klepsydra AI allows integration of custom mathematical backends on various hardware, enabling mission-specific AI deployment.
- **Middleware Support:** Native support for SpaceWire, CSP, ZeroMQ, and ROS2 provides a plug-and-play foundation for space-ready AI services.
- **Performance and Adaptability:** Klepsydra AI delivers high-performance and adaptable AI services for space systems.



Parallel data streams maximize GPU utilization with minimal power impact and seamless integration—unlocking the ability to run more Edge AI algorithms on GPUs.

Klepsydra GPU Connector

- **Parallel GPU Execution:** Klepsydra enables efficient parallel execution of CUDA kernels, CUDA Graphs, and TensorRT workloads on GPUs.
- **Low Latency and High Occupancy:** Zero-copy memory handling maximises GPU occupancy while maintaining ultra-low latency.
- **Concurrent Processing:** Allows for concurrent execution of pre-processing, AI inference, and post-processing on the same GPU, leading to faster and more efficient performance.



	NVIDIA Tools	Klepsydra GPU Connector
Inference engine	TensorRT	GPU Connector (CuDNN-based) & TensorRT
GPU Occupancy	Medium	High
Cyber-security	No	Yes
Safety Certification	No	Yes For the CPU part)
Secure OS Support	No	VxWorks, DDC-i (For the CPU part)



Fast, secure and safe AI software for Edge AI Systems



Software is available at
<https://downloads.klepsydra.com/>
Or contact
sales@klepsydra.com

Available in three formats:

- Docker images
- Debian packages
- Tarballs

A trial license covers:

- A 90-day license
- 10 hours of technical support
- A 1-hour technical workshop

Contact

Klepsydra Technologies AG
6300 Zug, Switzerland

www.klepsydra.com
sales@klepsydra.com

 +Klepsydra Technologies
 wa.me/41782493720