EDGECORTIX®

SAKURA-II M.2 Module
Datasheet

[Revision History](#)

TABLE OF CONTENTS

SAKURA-II Edge AI M.2 Modules Datasheet v0.71 - PRELIMINARY

1. OVERVIEW

This User Manual covers the EdgeCortix SAKURA-II M.2 Modules. These evaluation platforms are ready to drop into an Ubuntu host PC slot for software development and AI model evaluation tasks. The SAKURA-II M.2 Modules feature an EdgeCortix SAKURA-II chip, an edge AI accelerator that is run-time reconfigurable using the EdgeCortix MERA compiler and software framework, and that boasts 60 TOPS using EdgeCortix's Dynamic Neural Accelerator (DNA).

ORDERING INFORMATION

Module Type	Ordering Part Number	Description
M.2 2280 Module	SAKURA-II M.2 Module Part Number TBD	Single SAKURA-II device, 16GB of DRAM and associated circuitry

2. M.2 MODULE FEATURES

Specification	M.2 Module
AI Accelerator	Single SAKURA-II
Performance	60 TOPS (INT8) 30 TFLOPS (BF16)
LPDDR4 DRAM	16GB (2 banks of 8GB)
PCIe Interface	Gen 3.0 x4
Board Management Controller (BMC)	Power sequencing, configuration, and reset Voltage, current, and temperature monitoring Protection shut-down SPI Interface to SAKURA-II device
USB Interface	USB-C connector that provides access to BMC for monitoring and control
Cooling Options	Fan Sink (12V)
Power Consumption	10W (typical)
Form Factor	2280-D6-M

4. M.2 MODULE DIMENSIONS (without fansink)

Specification	M.2 Module
PCB Length	80.00 +/- 0.15 mm
PCB Width	22.00 +/- 0.15 mm
PCB Thickness	0.80 +/- 0.08 mm
Top-Side Component Height	3.2 mm max
Bottom-Side Component Height	1.5 mm max

4.1 M.2 MODULE DIMENSIONS (with fansink)

Specification	M.2 Module
Length	80 mm
Width	24.5 mm
Height	28.5 mm

5. M.2 MODULE IMAGE

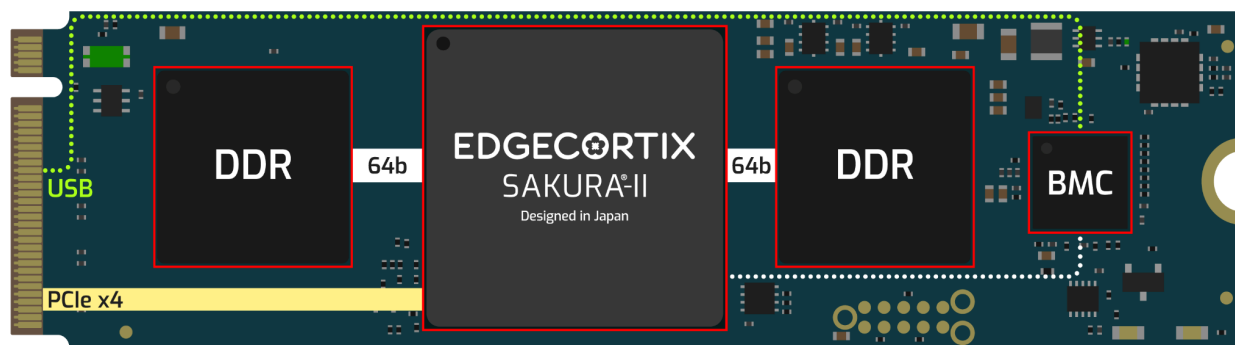Figure 1 shows an annotated image of the SAKURA-II M.2 Module, with important components identified:

FIGURE 1: SAKURA-II M.2 Module Annotated Image

Figure below shows an image of the SAKURA-II M.2 Module with the fansink integrated and it includes a 3-pin connector for powering the fan. This fan must be connected and running (i.e powered) to ensure proper cooling of the M.2 module to avoid overheating and potential thermal shutdown., with important components identified:

The fan connector is keyed and fits into a standard 3 or 4-pin fan connector in most PC systems. If your system does not have a way to connect and power the fan, see below for details on the connector and pinout. The fan runs using 12V and below are the connector pinout and recommended connections when using either a 3-pin or 4-pin connector/header.
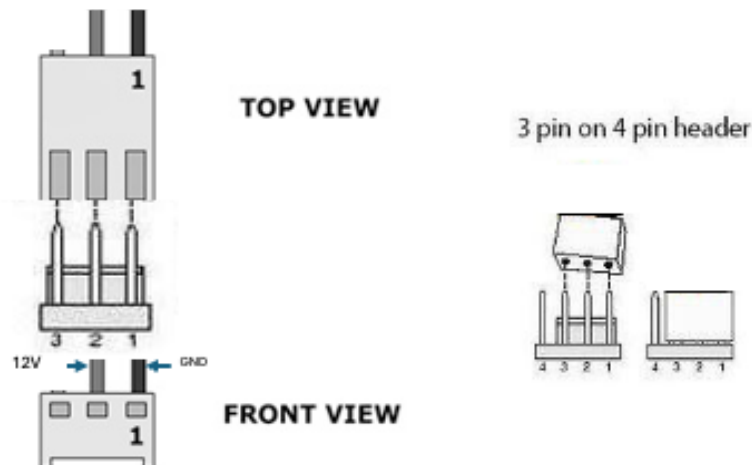
If you have questions about thermal considerations, please contact EdgeCortix.

6. M.2 MODULE APPLICATIONS

The SAKURA-II PCIe Cards are intended for use in a wide range of applications across a variety of market sectors. Primarily targeted at next generation AI functionality integrated into new and existing systems, the SAKURA-II PCIe Cards can be easily added to any system with a PCIe backplane.

Efficient Edge AI processing can provide many additional features, including:
- Natural language processing
- Object/person recognition
- Segmentation and Identification
- Obstacle Identification and Avoidance
- Advanced Signal Processing

Some of the markets that SAKURA-II is well suited for:

Smart City: For critical city management applications like traffic control, image recognition and identification, SAKURA-II can operate cost-effectively throughout the city environment, in applications like traffic control, image recognition and identification, SAKURA-II can operate cost-effectively throughout the city environment.

Smart Retail: Retailers can use SAKURA-II at the edge for inventory management, personalized recommendations, customer behavior analysis, smart shelving, checkout systems, and supply chain optimization.

Smart Appliances: Using SAKURA-II to manage appliances and choose ideal operation in real time can lead to significant energy use reduction and cost savings using AI.

Smart Manufacturing: In factories, edge AI solutions optimize production lines, predict equipment failures, and enhance quality control. Real-time analysis of sensor data helps improve efficiency and reduce downtime (and product quality) using SAKURA-II.

Smart Agriculture: Using SAKURA-II in the fields to manage irrigation and fertilization usage can increase yields and reduce costs with real-time soil and weather analysis.

Security: Surveillance cameras equipped with SAKURA-II can detect anomalies, recognize faces, and identify potential threats. This technology enhances public safety and reduces response time.

Autonomous Vehicles: Edge AI decision making is critical for human safety and SAKURA-II can use real-time processing to support accurate analysis and decision making for safe vehicle operation.

Robotics: Robots and autonomous machines rely on SAKURA-II capabilities for real-time perception, navigation, and decision-making. This is crucial for applications like warehouse automation and delivery drones.

Space/Aviation: SAKURA-II can assist in aircraft maintenance, engine performance, component wear prediction and improved safety and reliability. SAKURA-II's radiation survivability allows operation in Space environments.

AI-RAN and 5G/MEC (Multi-access Edge Computing): SAKURA-II can be leveraged to improve network management, predict network congestion, and enhance user experience, and can enable faster decision-making in complex telecom environments – and delivers new capabilities in the 5G/MEC space.

7.   M.2 MODULE SPECIFICATIONS

7.1   ABSOLUTE MAXIMUM RATINGS

Stresses outside the AMR may cause permanent damage; extended operation at AMR may degrade performance and affect reliability.

ABSOLUTE MAXIMUM RATINGS	Minimum	Maximum	Units
Storage Temperature	-40	+85	ºC
Operating Temperature	-20	+85	ºC
3.3V Power Supply	3.00	3.60	V
12V Power Supply	11.04	12.96	V

7.2   ELECTRICAL AND ENVIRONMENTAL SPECIFICATIONS

All electrical specifications for the M.2 Modules comply with the M.2 specification, which can be found on the PCI-SIG website at Specifications | PCI-SIG..

Specification	M.2 Modules
Electrical	Onboard power derived from M.2 slot (3.3V)
Environmental	-20C to +85C (component operating range) 0 to 95% humidity (non-condensing)

○   ## 7.3   PEAK PERFORMANCE

Peak Performance	M.2 Module
Trillion Operations per Second (TOPS) using INT8	60
Tera Floating-point Operations per Second (TFLOPS) using BF16	30

7.4 POWER CONSUMPTION

The power consumed by the M.2 Modules can vary depending on the model used, the number of inferences per second, and the operating frequency being run. The different M.2 Modules will have different power profiles due to the Compute Blocks used, size of memory and other factors.

7.5 POWER AND THERMAL MANAGEMENT

Users must consider the effects of the operation of the M.2 Modules on the system it is inserted into. Peak currents must be handled properly by the host system so as to not cause damage to other boards or components in the system. Currents are highly variable depending on the models used and spikes can occur. Your host system must be able to tolerate these higher currents, and your power supply must provide fast transient response performance.

The M.2 Modules consume power and the heat generated must be actively handled to ensure excessive board and component temperatures are avoided. If you have a system with active cooling sufficient to keep the heat down, you can select the heat sink option. If the system does not have active cooling, it is recommended that you choose the fan sink option to ensure the proper cooling is achieved.

7.6 ESD PROTECTION

SAKURA-II M.2 Modules are populated with electrostatic discharge (ESD) sensitive devices which can be damaged by static charges that can build up on people, tools, and other surfaces. Proper care must be taken in handling these devices and proper grounding must be maintained to ensure that any ESD does not damage any devices on the M.2 Module. It is beyond the scope of this document to explain and provide specific ESD protection schemes, but users should be familiar with these processes that apply to all ESD-sensitive semiconductor devices. No warranty is provided for improper handling of the SAKURA-II M.2 Module and damage to any devices on the Card is the full responsibility of the user.

8.   M.2 MODULE CONNECTOR PINOUT

All connector pinout specifications for the M.2 Modules comply with the M.2 specification, which can be found on the PCI-SIG website at <u>Specifications | PCI-SIG</u>..

9.   TOOLS, DOCUMENTATION, AND SUPPORT

9.1   INCLUDED COMPONENTS

This evaluation kit includes the following components:
- SAKURA-II M.2 Module
- Associated software:
  - MERA Compiler and Software Framework
    - Refer to the [MERA Installation and User Manual](#) for instructions
  - Installation script to automatically prepare a Python virtual environment suitable to run all the provided scripts and demos
  - Demos showcasing how to get detections from models that perform monocular depth estimation and object detection. A demo of how to take two models and fuse them into a single model, compile them, and get detections from hardware and use them to benchmark this model.
  - NOTE: All items are available in the EdgeCortix Developer Zone.

9.2   PRODUCT CERTIFICATIONS AND SPECIFICATIONS

Category	M.2 Modules
Product Certifications	FCC, CE, UKCA Japanese/Korean versions UL
Manufacturing Specifications	RoHS IPC-A-610 Class 2

9.3 HOST AND SOFTWARE REQUIREMENTS

Host Platform	x86-64 or ARM
Operating System	Ubuntu 22.04 LTS OR 24.04 LTS
Development Platform	MERA-II Compiler Framework including the MERA Quantizer
ML Frameworks	PyTorch, ONNX, TensorFlow Lite
Compiler Framework	Apache TVM MLIR
Models	Source from Hugging Face or EdgeCortix Model Library

* For compiling and generating run-time files (for deployment), users should use x86 systems only. Run-time files can be generated for both x86 or Arm architectures. Arm systems can then be used to deploy the Arm architecture generated runtime files.

9.4 OTHER DOCUMENTATION

In addition to this datasheet, other available documentation includes:
- SAKURA-II M.2 Module User Manual
  - Includes installation and usage details for the M.2 Modules
- SAKURA-II Device Datasheet
  - Provides detailed information on the SAKURA-II device

9.5 SOFTWARE DOWNLOAD AND TECHNICAL SUPPORT

To get access to the MERA Compiler framework and get technical support from EdgeCortix, users should first join the Developer Zone. You can put in a request for access here. In the Developer Zone, you will find:
1. MERA Compiler Software downloads
2. MERA Tutorials
3. Documentation listed above in Section 9.4
4. Ticket System for technical support

10.   REVISION HISTORY

Revision	Date	Summary
0.71	March 2025	Updated M.2 PCB dimensions in section 4
0.7	February 2025	Initial Release