

Hailo-8

M.2 AI Acceleration Module

(NGFF M.2 2242/2260/2280 Key M Card)



Data Sheet
HM218B1C2FA

Rev 0.7
March 2021

Disclaimer

Copyright

© 2021 Hailo Technologies Ltd ("Hailo"). All Rights Reserved.

No part of this document may be reproduced or transmitted in any form without the express, written permission of Hailo. Nothing contained in this document should be construed as granting any license or right to use proprietary information without the written permission of Hailo.

This version of the document supersedes all previous versions.

General Notice

Hailo, to the fullest extent permitted by law, provides this document "as-is" and disclaims all warranties, either express or implied, statutory or otherwise, including but not limited to the implied warranties of merchantability, non-infringement of third parties' rights, and fitness for particular purpose.

Hailo assumes no liability for any error in this document and for damages, whether direct, indirect, incidental, consequential or otherwise, that may result from such errors, including but not limited to loss of data or profits.

The content in this document is subject to change without prior notice. Hailo reserves the right to make changes to said content without prior notification to users.

Documentation Control

Revision History

Version	Date	Description
0.1	July 2020	First draft
0.2	July 2020	Early access version – approved for limited disclosure
0.3	August 2020	Format improvements, updated diagram and tables
0.4	September 2020	Minor updates and corrections
0.5	December 2020	Minor updates and corrections
0.6	February 2021	Added information on power consumption, updated installation instructions
0.7	March 2021	Changes to formatting and phrasing

Table of Contents

1. Overview	5
1.1. Introduction	5
1.2. System Requirements	5
1.3. System Level Block Diagram	6
2. Specifications	7
2.1. Key Properties	7
2.2. PCIe Connector Pinout	8
2.3. Mechanical Properties	9
3. Power Consumption and Thermal Management.....	10
3.1. Module Power Consumption	10
3.2. Thermal Management Requirements.....	10
3.3. Overheat and Overcurrent Protection.....	11
4. Installation and Troubleshooting.....	12
4.1. Hardware Installation	12
4.2. Software Installation.....	12
4.3. Troubleshooting.....	14

1. Overview

1.1. Introduction

The Hailo-8 M.2 AI Acceleration Module (part number HM218B1C2FA), compatible with the NGFF/M.2 form factor, is targeted at artificial intelligence (AI) applications. It is a member of Hailo's comprehensive family of PCI Express (PCIe) based acceleration modules that meet industry standards for a range of form factors and performance objectives.

The module is based on the Hailo-8 AI processor and features a full PCIe Gen 3 4-lane interface (x4), which enables high throughput of input and output data.

As a PCIe device, the module can be used to perform real-time, low latency neural network inference, using PCIe for streaming input data and for streaming inference results.

1.2. System Requirements

- **Hardware:**

- CPU Architecture
 - x86 based; or
 - ARM aarch-64 based
- At least 1 available M.2 slot (M keyed, Socket 3 type, sometimes called NVME)

- **Software:**

- Linux Ubuntu or other distributions
- Supported kernel versions: multiple, tested on Linux kernel versions 4.15.0-39-generic and 5.0.16-050016-generic.

1.3. System Level Block Diagram

The following block diagram illustrates the Hailo-8 M.2 AI Acceleration Module (HM218B1C2FA):

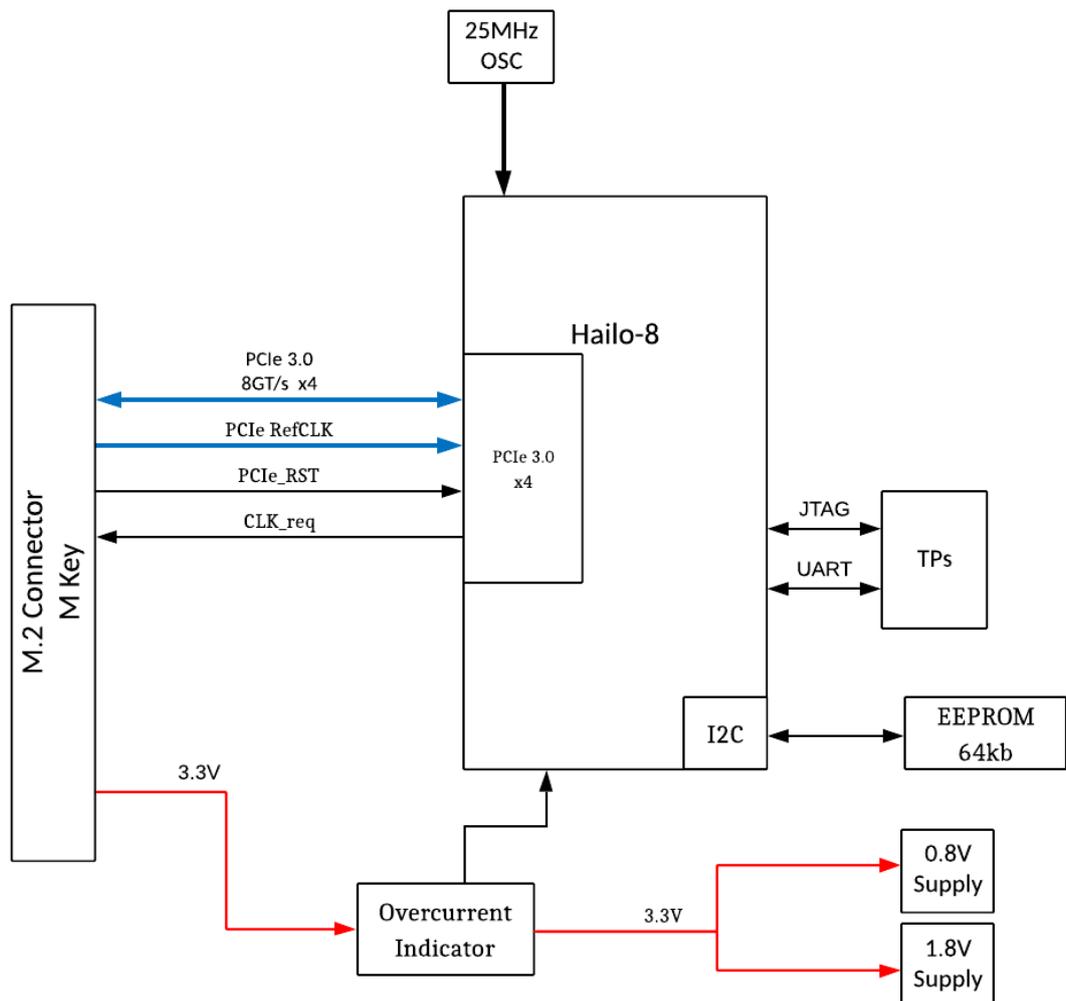


Figure 1: HM218B1C2FA Functional Block Diagram

2. Specifications

2.1. Key Properties

Compliance	
Certification	CE; FCC Class A
Environmental	
Storage Temperature	0°-70° C
Operating Temperature	0°-70° C (<i>refer to section 3.2 for further details</i>)
Physical	
Form Factor	NGFF M.2 Key M
Dimensions	22 x 42 / 22 x 60 / 22 x 80 mm
Electrical	
Power supply	3.3V ± 5%
Thermal design power (TDP)	8.65W
Interface	PCIe Gen 3, 4-lanes (x4)

2.2. PCIe Connector Pinout

Pin Number(s)	Signal	Type
21	CONFIG_0 (GND)	Defines module type: SSD-PCIe
69	CONFIG_1 (NC)	
75	CONFIG_2 (GND)	
1	CONFIG_3 (GND)	
3,27,33,39,45,51,57,71,73	GND	Power
2,4,72,74	3.3V	Power
50	PERST#	I
52	CLKREQ#	I/O
54	PEWAKE#	I/O
53,55	REFCLK[n/p]	I
41,43	PET[n/p]0	O
47,49	PER[n/p]0	I
29,31	PET[n/p]1	O
35,37	PER[n/p]1	I
17,19	PET[n/p]2	O
23,25	PER[n/p]2	I
5,7	PET[n/p]3	O
11,13	PER[n/p]3	I

Table 1: HM218B1C2FA Pin Description

For more information, see the *PCI Express M.2 Specification Revision 3.0* on

the [PCISIG website](#).

2.3. Mechanical Properties

Below is a mechanical outline for the HM218B1C2FA module. Note that even though the module is a small form factor 22mm*42mm (2242), it is provided with removable tabs to accommodate systems which only support 2260 or 2280 modules mechanically.

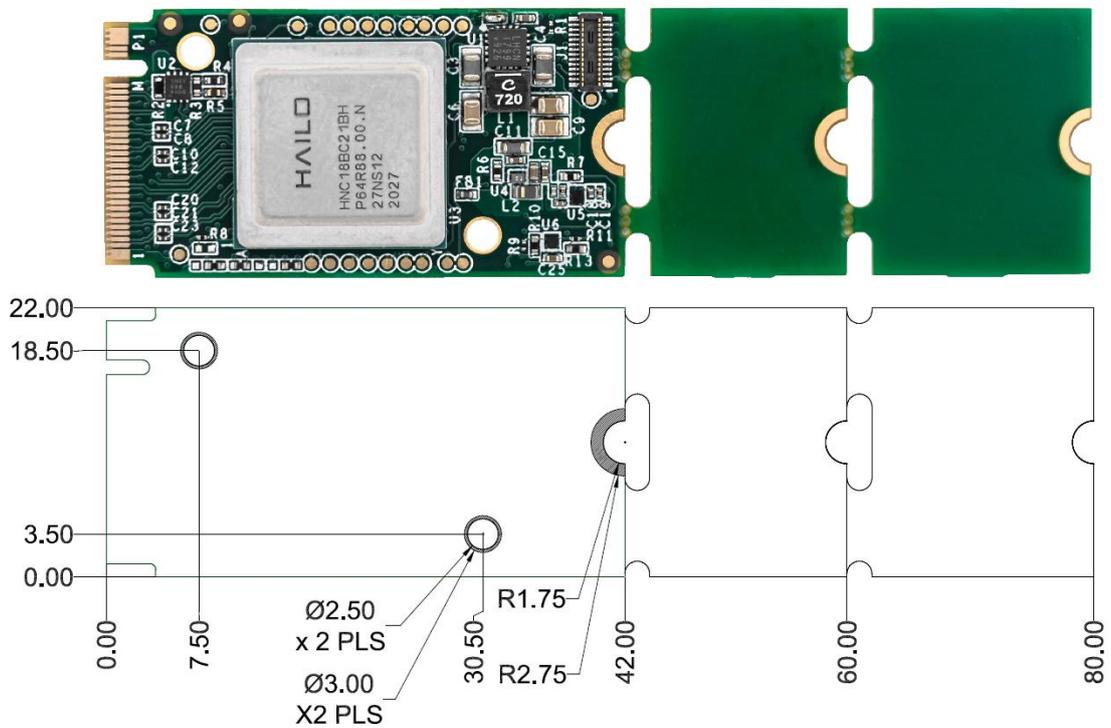


Figure 2: HM218B1C2FA Mechanical Outline (Top View)

For more information, see the *PCI Express M.2 Specification Revision 3.0, Version 1.2* on the [PCISIG website](#).

3. Power Consumption and Thermal Management

3.1. Module Power Consumption

The module requires a 3.3V supply (power pins are as detailed in Table 1).

The module’s power consumption is dependent on the resources utilized for inference. Maximum power consumption is 8.25W (or 2.5A total current draw from 3.3V pins) at full utilization.

The module’s power consumption (typical at 25° C) is listed below for typical configurations:

Configuration	Power [W]
Resnet-50 224x224 @ 915 FPS	3.3
MobileNet-SSD 300x300 @ 996 FPS	2.4

Table 2 - Power Consumption for Typical Configurations

For up-to-date benchmark performance please visit the benchmarks page on Hailo’s website.

3.2. Thermal Management Requirements

The HM218B1C2FA is tested and validated for commercial temperature grade (0°C – 70°C) when mounted in a test fixture.

Proper heat dissipation must be employed to ensure that the Hailo-8 chip does not overheat. The Hailo-8 is designed to dissipate most of the heat to the top surface of the package.

For more information on typical use configurations and design guidelines, see Hailo's *Hailo-8 AI Acceleration Module Thermal Design Considerations Application Note*.

3.3. Overheat and Overcurrent Protection

The module features monitoring and protection from overheat conditions and excessive power consumption. The following features are supported:

TBD

For additional information on thermal design, see sec. 3.2.

For more information on power consumption limits, see the *PCI Express M.2 Specification Revision 3.0* on the [PCISIG website](#).

4. Installation and Troubleshooting

4.1. Hardware Installation

1. Make sure that the host computer is turned off.
2. Firmly attach the M.2 module into a proper M.2 slot (make sure it is a Socket 3 type suitable for M key cards).
3. Power up the computer and log on to Linux.
4. Verify that the Hailo-8 is detected on the PCIe bus. From a Linux terminal, type:

```
>> lspci | grep "Co-processor: Hailo Technologies Ltd. Hailo-8 AI
Processor (rev 01)"
```

4.2. Software Installation¹

1. Power up the computer and log on to Linux.

Locate the device driver installation package and extract the folder *Hailo-8 PCIe driver* from the .tar file.

Install driver. From a Linux terminal, type:

```
>> chmod u+x install.sh
>> ./install.sh
```

Verify the device is identified properly using the Hailo-RT API.

- a. Start the Hailo-RT virtual environment. From a Linux terminal type, at the location where the Hailo-RT package was installed:

¹ All software related guidelines are based on pre-released software and may change upon product release. Any such changes will be included in future versions of this document.

```
>> source hailo_virtualenv/bin/activate
>> hailo fw-control scan
```

Run python² and type the following in the python interpreter to identify the device:

```
>> from hailo_platform.drivers.hw_object import HailoPcieObject
>> p = HailoPcieObject()
>> p.control.identify()
```

For further information and instructions, see *Running HailoRT installation* in the HailoRT User Guide.

² The method proposed is based on the Python API. Usage of python is optional and provided here as an example.

4.3. Troubleshooting

A. Improper PCIe device enumeration	
How to verify:	<p>From a Linux terminal type:</p> <pre>>> lspci grep "Co-processor: Hailo Technologies Ltd. Hailo-8 AI Processor (rev 01)"</pre> <p>The device should be listed in the terminal output</p>
Possible root cause:	Improper mechanical installation
Solution:	Verify the module is properly attached and secured into the M.2 slot.
Possible root cause:	Slot is not functional
Solution:	<p>Verify the slot in use is a valid M.2 M key slot.</p> <p>Check to see if the slot is disabled in platform BIOS.</p>

B. Device driver not properly installed	
How to verify:	<p>From a Linux terminal type:</p> <pre>>> lsmod grep <TBD></pre>

	The device should be listed in the terminal output
Possible root cause:	Driver not installed
Solution:	Obtain the driver installation package provided by Hailo and follows SW installation instructions (section 4.2)

C. Module not identified by Hailo-RT	
How to verify:	<p>From a python shell type:</p> <pre>>> from hailo_platform.drivers.hw_object import HailoPcieObject >> p = HailoPcieObject() >> p.control.identify()</pre> <p>The device should be listed in the interpreter output</p>
Possible root cause:	Improper installation of the Hailo-RT package
Solution:	Re-install the Hailo-RT package by following the installation instructions.