D2.2: Specifications and architecture design

Document Properties

Contract Number	**101189612**	
Contractual Deadline	M6 (30th of June, 2025)	
Dissemination Level	Public	
Nature	Report	
Edited by :	Sven Schenk, Extoll	
Authors	Olivier Déprez, SiPearl Blagovest Tushev, CloudSigma Xavier Teruel, BSC Thomas Van Liefde, 2CRSi Çağatay Yilmaz, RISE	Dimitris Theodoropoulos, ICCS Iakovos Mavroidis, Exapsys Fabien Chaix, FORTH Mariam El Hassouni, SiPearl Sven Schenk, Extoll
Reviewers	Ivy Peng, KTH, Manolis Marazakis, FORTH, Josep Sans, SMD	
Date	22/07/2025	
Keywords	RISC-V, Cloud, Acceleration, Network Object Store, Containers, PCIe, CXL, IaaS, PaaS, OCP, DC-MHS, HPM, DC-SCM, BMC, Memory disaggregation	
Status	Final	
Release	1.0	

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no101189612. The project is funded under the call on "Digital and emerging technologies for competitiveness and fit for the green deal"

History of Changes

Release	Date	Author, Organization	Description of Changes
0.1	22/05/2025	Sven Schenk, Extoll	Definition of ToC
0.2	05/06/2025	Sven Schenk, Extoll	Initial input to section 3 & 5
0.3	11/05/2025	Sven Schenk, Extoll, All	Added input to section 3
0.4	16/05/2025	Sven Schenk, Extoll, All	Updated input to section 3
0.5	22/05/2025	Sven Schenk, Extoll, All	Updated input to section 5
0.6	29/06/2025	Sven Schenk, Extoll	Input to sections 2 & 4
0.7	08/07/2025	Sven Schenk, Extoll, All	Updated inputs to sections 3,4,5
0.8	12/07/2025	Ivy Peng, KTH, Manolis Marazakis, FORTH, Josep Sans, SMD	Document ready for review
0.9	18/072025	Sven Schenk, Extoll, All	Addressing reviewers comments
1.0	22/07/2025	Sven Schenk, Extoll	Final version submitted

Table of Contents

List of Figures

List of Tables

1 Executive Summary

Building on top of work carried out in previous tasks of this WP, this document gives specifications of the hardware and software that are required to fulfil the requirements defined in [D2.1]: "Requirements and use cases refinement". While adhering to the Open Compute Project (OCP) server family standards, more details are given on the individual components. This includes the modules for computation and acceleration, alongside a system security/control module, all operating with fully-featured operating systems and runtimes in OCP compliant servers.

To achieve this, detailed specifications on each component will be derived from the high level requirements that have been developed in T2.1. With these specifications, this document serves as a reference for the technical work carried out in the WPs developing the components. Also, a make or buy analysis for the security and control module, as well as for the OCP-NIC is included to guide later WPs in their decision making.

We start by describing the DC-MHS server in general for SDNO and M-FLW use cases, followed by the specification of the HPM processor module for Rhea2 and then EPAC/EUPILOT based modules and also the management module with firmware and Platform Root of Trust component. These will be the guidelines for the implementation of the boards in WP3. Hereinafter, the boot firmware, the operating system, and the required device drivers are described, which will enable the anticipated use-scenarios. Finally, possible uses for the proposed servers in typical cloud infrastructures are discussed.

2 Context

2.1 Purpose of the document

This document represents the specifications for the HIGHER platforms aiming to meet the requirements defined in previous work of WP2 [D2.1].

Specifications are given for:

i. DC-MHS assembled servers.
ii. An OCP-compliant processor module dedicated to computation and hosting 2 RHEA2 EPI chips.
iii. An OCP-compliant processor module dedicated to acceleration and hosting EPAC 2.0 EPI chips or EUPilot chips.
iv. An OCP-compliant management module, hosting a RISC-V processor inside of an FPGA for server management, security, and control features.

A Make or Buy Analysis is also included to determine if the overall benefits of buying an existing management module surpass the flexibility of a self-developed module and furthermore, whether the same applies to the OCP-NIC.

This document is the second WP2 deliverable covering the work carried out in T2.2: "Specifications and architecture design" for the different parts of the servers to guide later WPs in their work of developing the respective parts. This document will inform T3.1 and T3.2 for developing the HPM processor modules, T3.3 for developing the DC-SCM module and Root Of Trust, and T3.4 for developing the server mechanics. Additional specifications are provided to these tasks to give further design considerations for the targeted use case of CXL memory disaggregation.

Furthermore, the associated software will be specified to inform T4.1 and T4.2 for developing the secure boot and OS for ARM and RISC-V HPMs respectively, and T4.3 for developing an open-source meta operating system that supports heterogeneous computing environments.

3 Specifications

3.1 DC-MHS Server

Two DC-MHS servers are going to be developed to cover as many use cases as possible. One server will be compatible with the M-SDNO specification from OCP while the second server will be compatible with the M-FLW specification from OCP. Having those two different servers will bring modularity to the systems and allow us to have as many architectures and possibility of interconnection between systems as possible. As of today, M-FLW is the most common OCP form factor for motherboards in dual CPU configuration. Both servers will be able to integrate boards with M-SDNO Class A form factor so they will be compatible with the EPAC/EUPILOT processor module developed for the HIGHER project. It has been decided to have a chassis compatible with M-FLW form factor and one compatible with M-SDNO Class C in case M-SDNO Class C isn't mostly adopted by OCP community members or motherboard manufacturers.

Having two different chassis allows a lot of flexibility and gives us insurance to be compatible with current or next generation of Host Processor Modules (HPMs). Both servers will have the same form factor, but internal architecture and compatibility will be different.

The server will be a standard 19" server. 19" is the most common dimension for servers in datacenters. The DC-MHS specification includes both 19" and 21". The focus will be on 19" at first, with a possibility to move easily to a 21" form factor in the future by adding some rails on the side of the server and modifying the way to power it up. The server will then be compatible with both 19" rack and OCP rack, allowing a good versatility for the server. The Width of a 19" server is around 482.6mm, including the ears on each side of the server, which allow the module to be fastened to the rack frame. The real available space for the hardware in terms of width is around 437mm (without ears and chassis).

For the height, the target is to have 2RU (Rack Unit), also referred as 2U. One U is around 44mm tall. To be able to integrate all the mandatory hardware within the server, at least a 2U size is required. The Thermal Design Power (TDP) of the different components of the server, like the CPU or the GPU, has been increasing in the recent times, needing more space to dissipate the generated heat. Hence, taller heat-sinks are needed to dissipate the higher TDP. The total height of the server will be 88mm. If it is not possible to have enough dissipation for the cooling of the different components, the servers would be 3RU height.

The depth aimed for is around 810mm, with the possibility to go up to 850mm if needed. We keep this possibility if there is a problem for cabling of the different components within the server or if we need to increase the thermal dissipation in the server depending on the hardware that is going to be used in the M-FLW server for example.

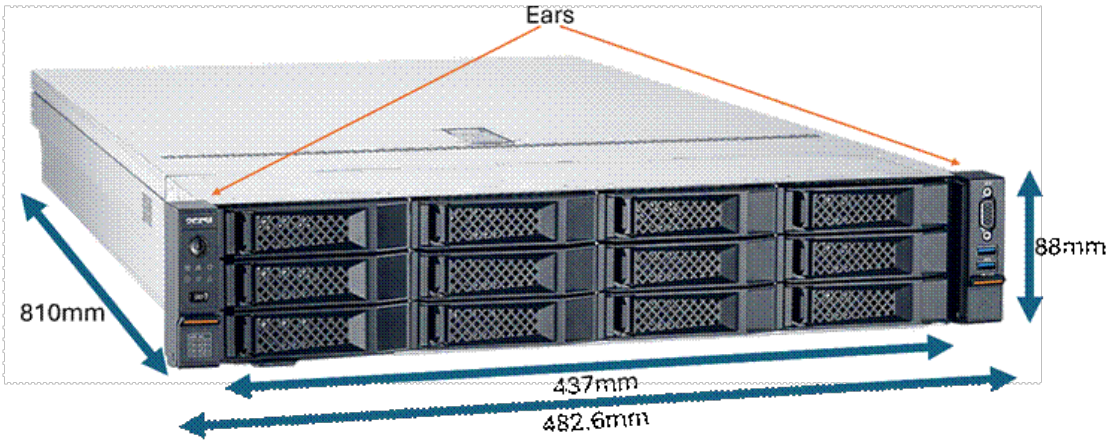

FIGURE 1 – SERVER DIMENSIONS

For future applications, to be compatible with the OCP form factor, which is 21", rails can be added to the side of the server for it to fit in a 21" bay. The way of powering up the server will also be modified to meet OCP requirement for [ORV3] (Open Rack version 3) specification.

3.1.1 M-FLW server

HPM compatibility

The server will integrate a motherboard with Full Width HPM Form Factor (M-FLW). This form factor is designed for 19" racks but is also compatible with 21" racks. M-FLW provides a lot of IO from the CPU, making it possible to interconnect a large number of devices within the server.

M-FLW is a standard which can be used with multiple generations of CPUs and memory. The chassis will be reusable with any type of CPU. It could be Intel, AMD or ARM-based CPUs, as long as the motherboard follows the M-FLW specification, especially for dimension and position of the different fixing holes.

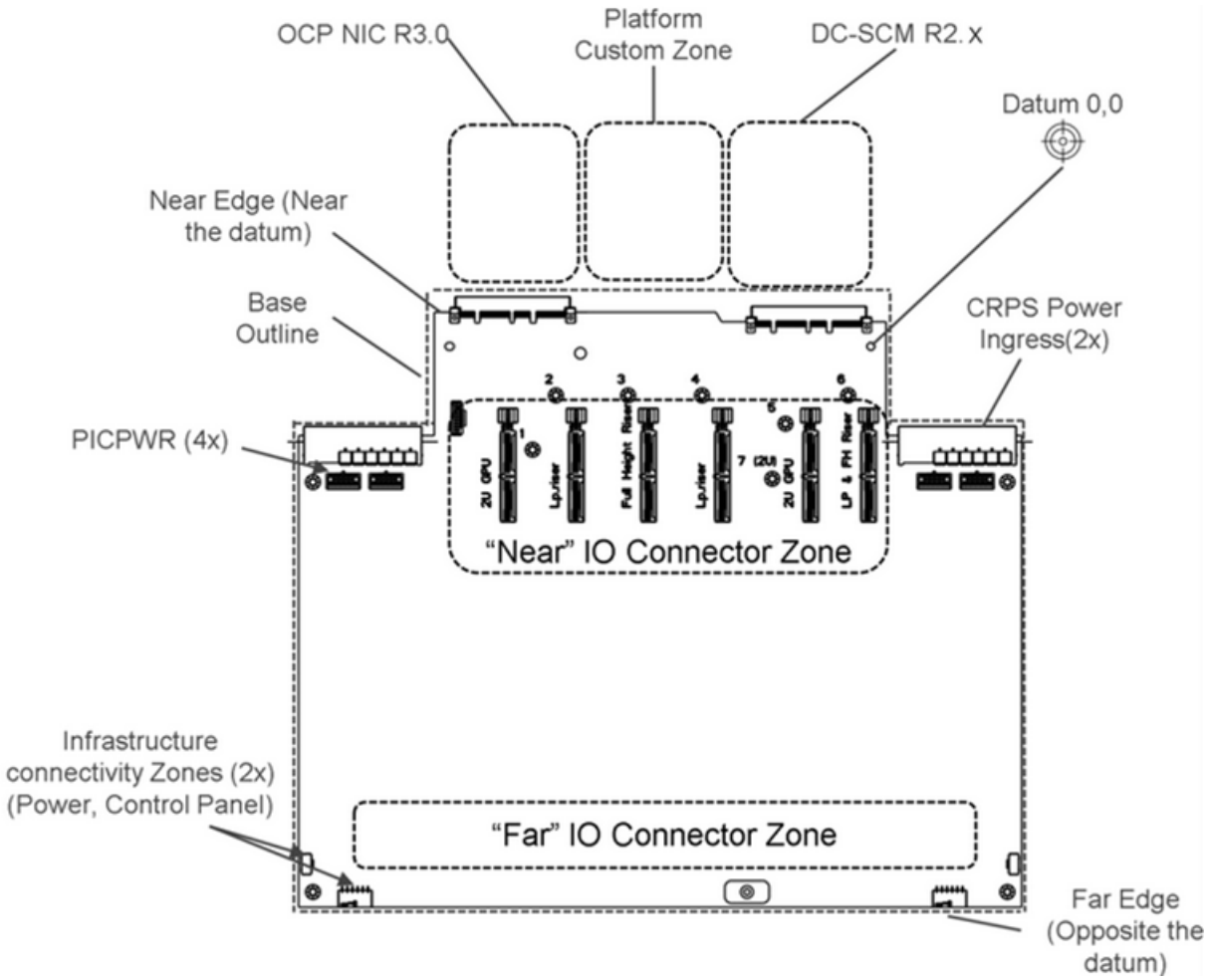

FIGURE 2 – M-FLW MOTHERBOARD OVERVIEW

The M-FLW specification allows multiple devices to be connected directly to the motherboard. At the front we have:

- One slot for an OCP NIC R3.0 card, generally used for network connectivity
- One slot for a DC-SCM R2.0 card. The DC-SCM is used for the management, control and security of the system. BMC is integrated in the DC-SCM module.
- One platform custom zone. Depending on the manufacturer, this zone can be either used for a second OCP NIC 3.0 slot or for storage (for example slots for E1.S drives). The use of this zone depends on the manufacturer of the M-FLW motherboard.

Two CRPS form factor power supply connectors are available on the motherboard. The server can be powered up directly via the motherboard but it is not mandatory. A system with a Power Distribution Board (PDB) can also power up the motherboard using the PICPWR connector available on the motherboard.

The Near IO Connector Zone purpose is to connect several types of devices to the motherboard. Up to 6 High Speed I/O (HSIO) connectors can be on the motherboard. Different types of connectivity are possible with HSIO connectors. It can be either a Board to Board (B2B) connection, a Board to Cable (B2C) connection or a Hybrid connection. Each connector is composed of one part for the power (up to 55A but mostly used with 21A limitation) and 2x separate 74 pins connections. Those connectors are SFF-TA-1033 connectors and are compliant with the M-XIO specification from OCP.

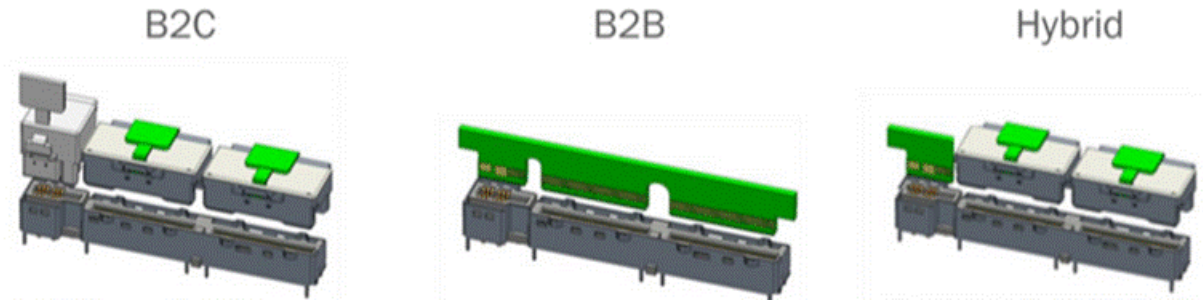

FIGURE 3 – HSIO CONNECTIVITY

Dimension and position of fixing holes and riser holes are a requirement for motherboard manufacturers who want their board to be compliant with the M-FLW specification from OCP. Several boards are already available on the market and more will be coming in the future.

In the figures below we can see some examples of M-FLW HPM available on the market today.

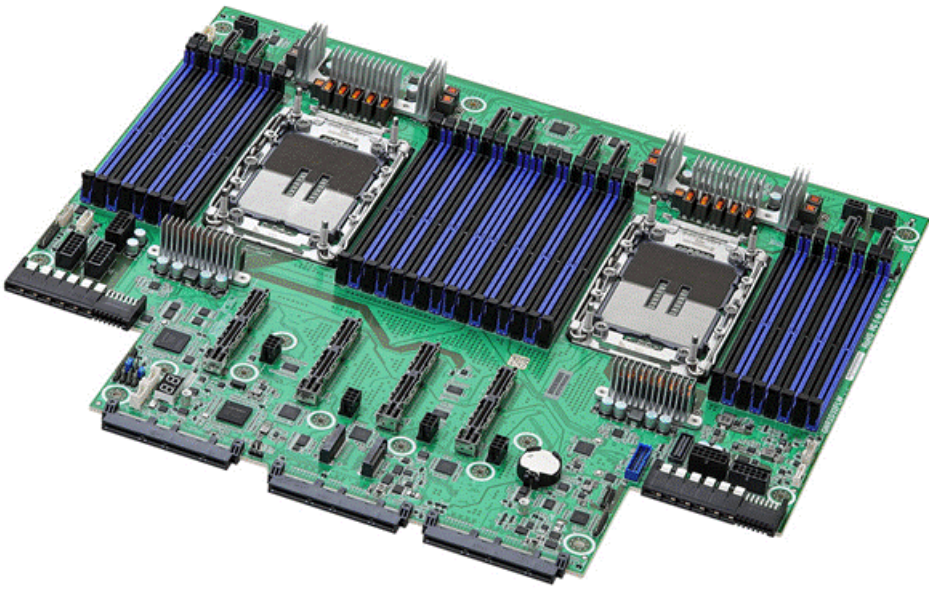

FIGURE 4 – GNR2D32FLW FROM ASROCKRACK

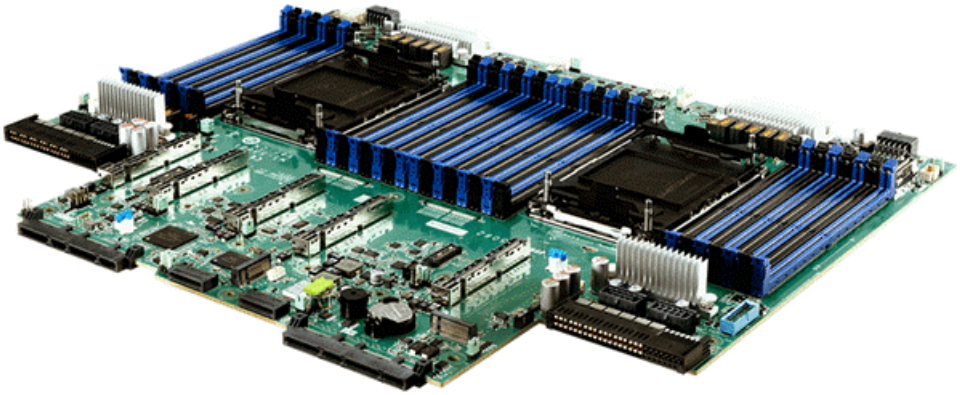

FIGURE 5 – D5062 FROM MSI

As we can see in figure 4, we have 2 slots for OCP NIC 3.0 card on ASrockRack motherboard while on figure 5 we have 2 slots for E1.S drives on MSI's one.

Power

The M-FLW HPM is powered by a 12V DC source. There are several options for powering up the system.

For 19" servers, power usually comes from a Power Supply Unit (PSU). Those PSUs are AC/DC and will provide the 12VDC to the system. Two M-CRPs connectors are on the motherboard, on which the PSUs are directly inserted. It is also possible to have the PSUs connected to a Power Distribution Board (PDB) allowing more flexibility in the positioning of the PSUs within the server. It will be possible to power up the different devices in the servers directly from the connectors on the PDB. M-PIC connectors are available on the M-FLW motherboard.

For the OCP standard, a powershelf, in which we have the PSUs, is connected to a busbar within the rack. Servers are connected to this busbar so there is no need to have a PSU inside the server, the power comes directly from the busbar. Following the latest OCP specification, [ORV3], the voltage is usually between 48V and 54V. Hence, a PDB is required to convert the 48V to the voltage needed in the server (usually 12V). Power management of the system and for the cooling solution (fans) can also be done on this PDB.

For the DC-MHS with M-FLW, we will have different options. PSUs will be either directly connected to the motherboard or they will be connected to a PDB at the back of servers. By having the space for the PSUs at the back, we can easily replace them by a cable connected to the busbar if we want to move from 19" to 21". There will be two PSUs in the server, each PSU will be up to 3200W, allowing a N+1 redundancy in some use cases, depending on the configuration of the server and TDP of the components (CPU, number of memories, GPUs, OAM, storage, etc..).

CRPS comes in different form factors. Most common is a width of 73.5mm and a depth of 185mm.

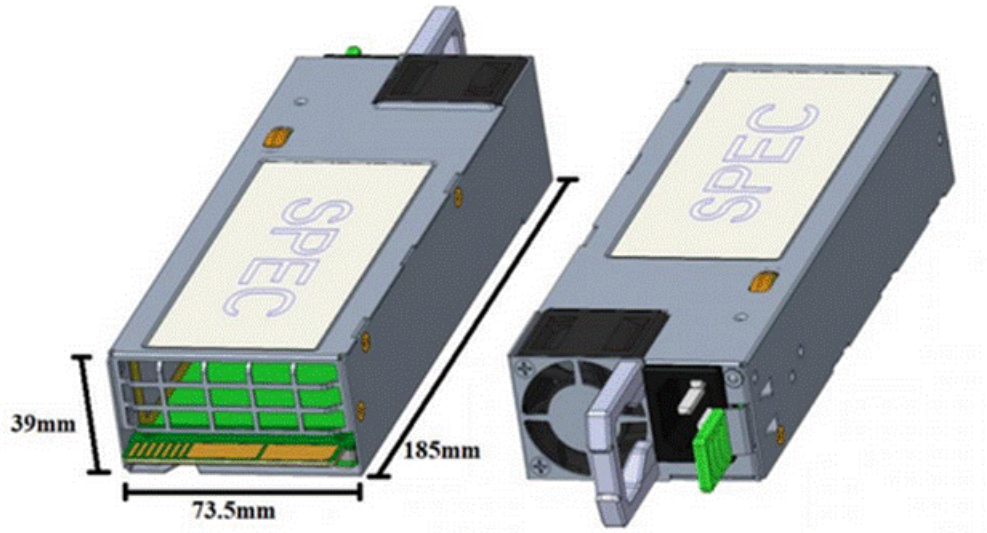

FIGURE 6 – CRPS FORM FACTOR

Both PSUs will be connected to one PDB. If necessary, a voltage convertor will be on the PDB in order to have every voltage needed to power up the different components of the server (5V and 3.3V if necessary). The PDB will have enough power connectors to cover every hardware option of the server. Connectors are mostly going to be M-PIC power connectors to be compliant with the OCP specification and to have an easy cabling solution for the power part.

Internal architecture and connectivity

The DC-MHS with M-FLW form factor motherboard will have a different configuration since it will be able to integrate different types of hardware.

As we saw previously, we will have two options for power supplies, either directly connected to the M-FLW motherboard or connected to a PDB and then we use cables to power up the different components of the server.

For the storage, it will be possible to have several drive form factors:

- E1.S: those drives can be directly connected to the motherboard, depending on the manufacturer design (MSI motherboard for example) or connected in an E1.S bay. Those bays will be half 5.25” width or 5.25” width. It will be possible to have up to 8x E1.S in a 5.25” bay. For PCIe gen 5, bays come with MCIO connectors and will then be connected directly to the motherboard (either to the MCIOs on the board or to the MXIO connectors). Storage can be either at the front or at the back (if the PSUs are connected to the PDB, there is empty space at the front where we can integrate some more storage).
- 2.5”: the drives will be inserted into a 5.25” drive bay (possibility to have also half 5.25” drive bay width). It is possible to connect up to four 2.5” drives in one 5.25" drive bay. The drives will be NVME SSDs in U.2 form factor with the possibility to have also U.3 and M.2 by using an adapter to convert from 2.5” to M.2. Depending on bay manufacturer, connectivity on the bay can be different. For PCIe gen 5 the standard should be MCIO connectors. Same as for the E1.S solution, the MCIO connectors on the bay will then be connected to the MCIO on the motherboard or MXIO through cables.
- M.2: there are usually M.2 slots available directly on the M-FLW motherboard (2280 and/or 22110 form factor). If more M.2 are needed, it will be possible to have an M.2 bay with same form factors as for E1.S and 2.5” available (half 5.25” bay drive or 5.25” bay drive). It is possible to use M.2 in a 2.5” bay drive with an adapter. Connectivity with the motherboard will be with MCIO cables directly connected to the MCIO on the motherboard or to the MXIO connectors.

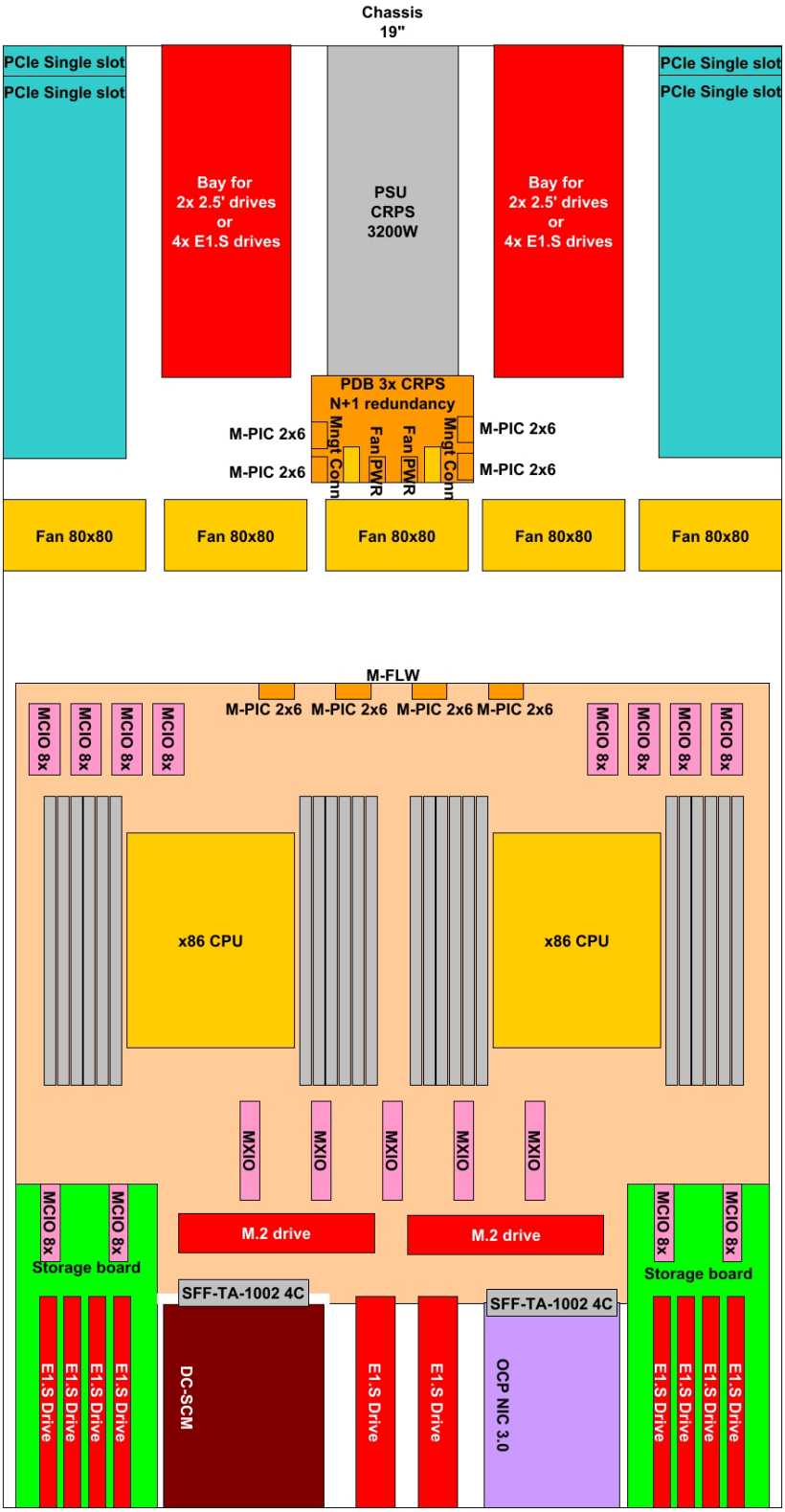

FIGURE 7 – M-FLW LOWER U OPTION 1

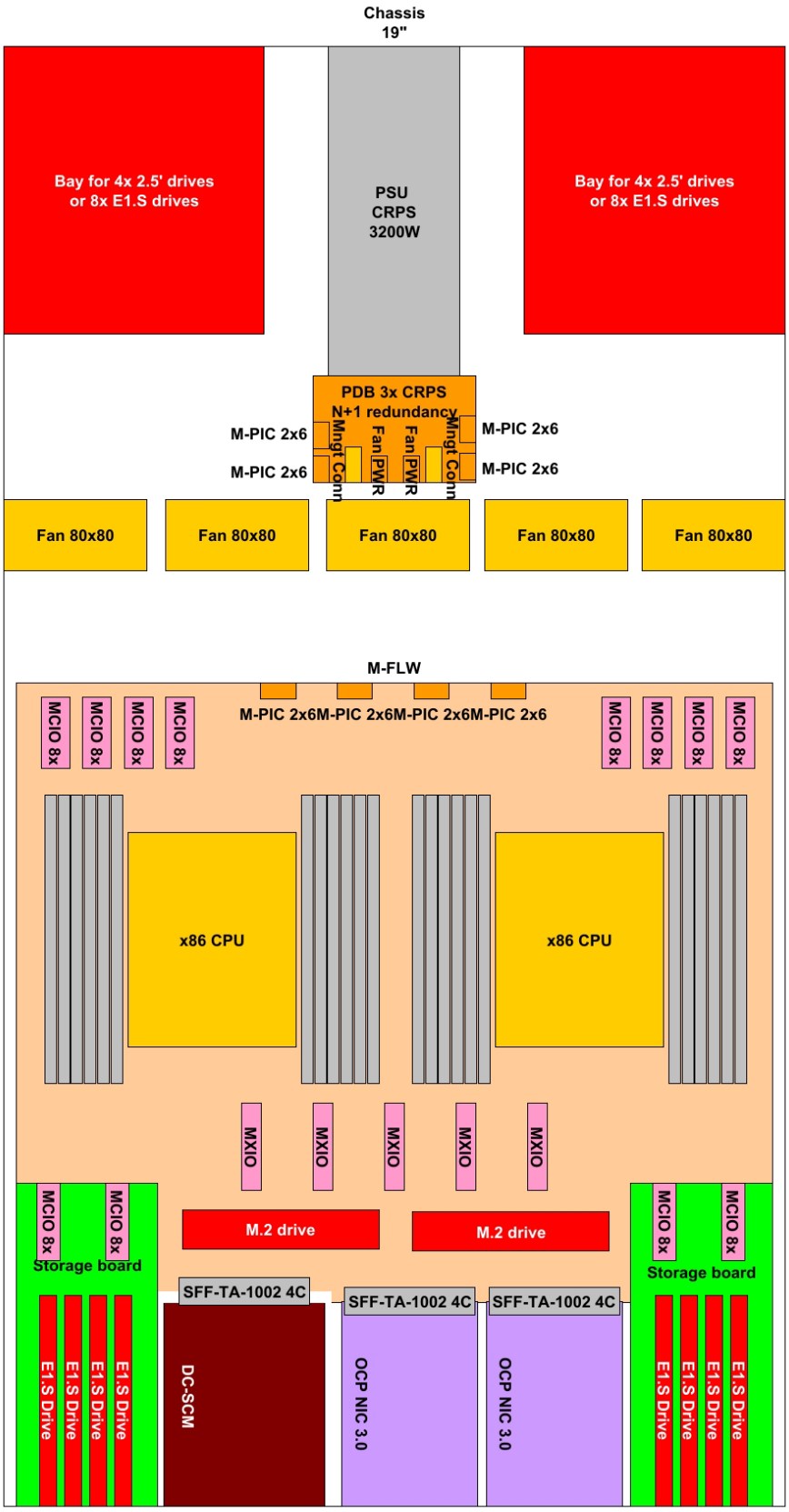

FIGURE 8 – M-FLW LOWER U OPTION 2

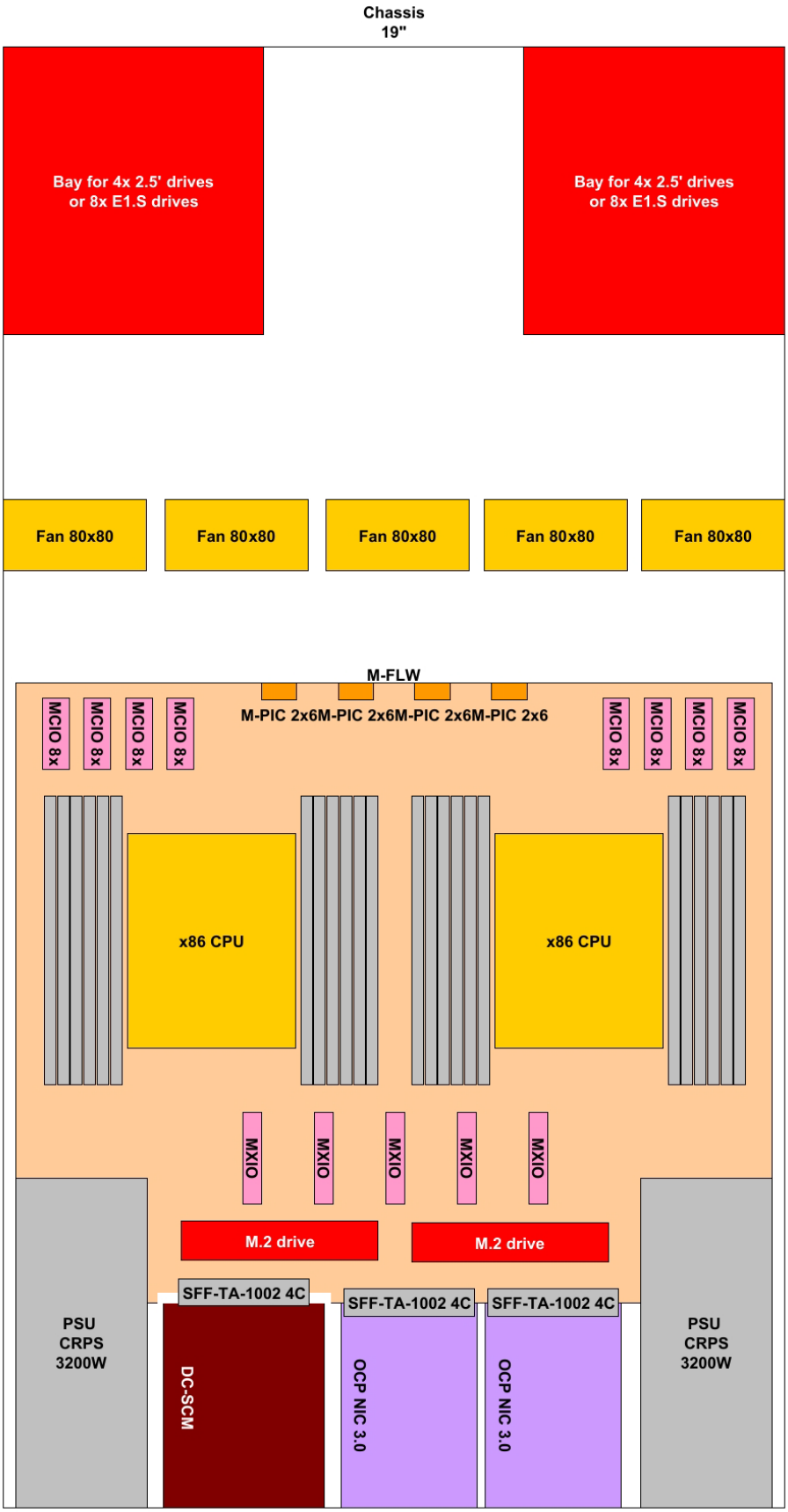

FIGURE 9 – M-FLW LOWER U OPTION 3

Figures 7 to 9 represent the lower U of the server. The server being 2U form factor, the first layer will be referred as lower U in this document, while the top layer, the second U, will be referred as upper U.

We have the M-FLW motherboard with two CPUs and the memory slots. For the moment, M-FLW motherboards on the market are available with x86 CPUs. M-PIC power connectors are available to power up the motherboard or to power up the different devices in the servers when CRPS are directly inserted in the motherboard. MCIO and MXIO connectors allow the connectivity of a lot of other devices in the server.

At the front of the motherboard, we have two solutions:

- 1x DC-SCM + 1x OCP NIC 3.0 + 2x E1.S drives
- 1x DC-SCM + 2x OCP NIC 3.0

We can see the different options for the power, with both CRPS PSUs being at the back of the server (one on top of each other, the second PSU is on the second U), connected to a PDB on which we have all the connectors required to power up every component of the server, or, CRPS PSUs directly inserted into the M-FLW motherboard. In case PSUs are connected to the PDB, the empty space at the front could be used by a storage board, allowing the connection of more storage drives. Depending on the availability and size of half 5.25" bay drives, a specific card might have to be developed in order to fit in the space of the CRPS power supply (73.5x185x39mm).

Storage options are represented with the different bays that can be integrated within the servers, giving the opportunity to have all the storage required and flexibility on type of storage depending on the use cases.

If half 5.25" bay drives are used, there is enough space to add PCIe slots at the back of the server. Those slots can be used to connect PCIe 16x (or less) devices to the server, like network card, RAID card, HBA card, etc... They can be connected to the MCIO or MXIO connectors on the motherboard via cables.

The server will be air cooled, with several fans of 80x80mm. Fans will be managed by the PDB and/or by the motherboard on the DC-SCM module connected to it. The BMC will reduce or increase fan speed depending on temperature and power consumption of components within the server.

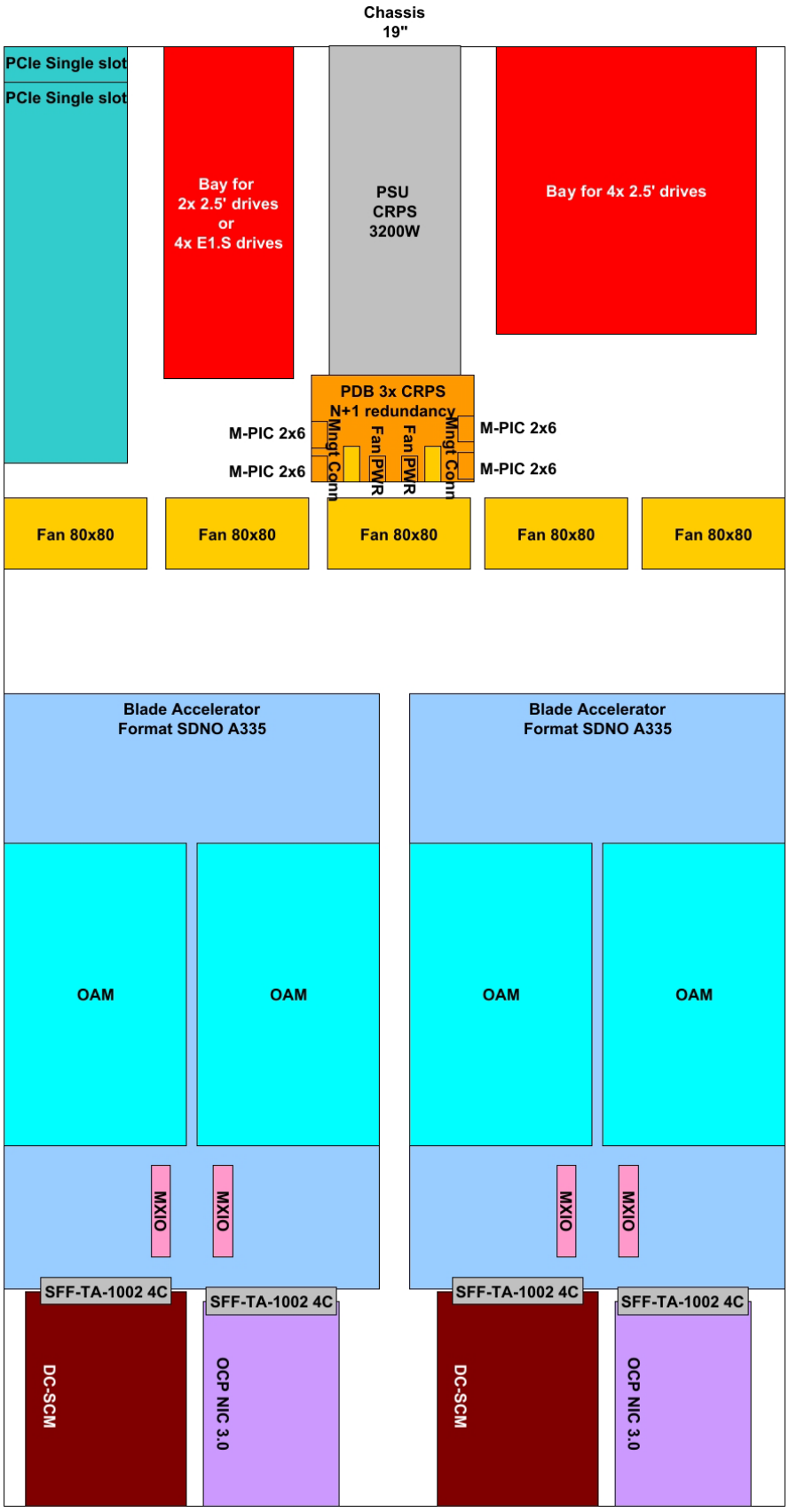

FIGURE 10 – M-FLW UPPER U OPTION 1

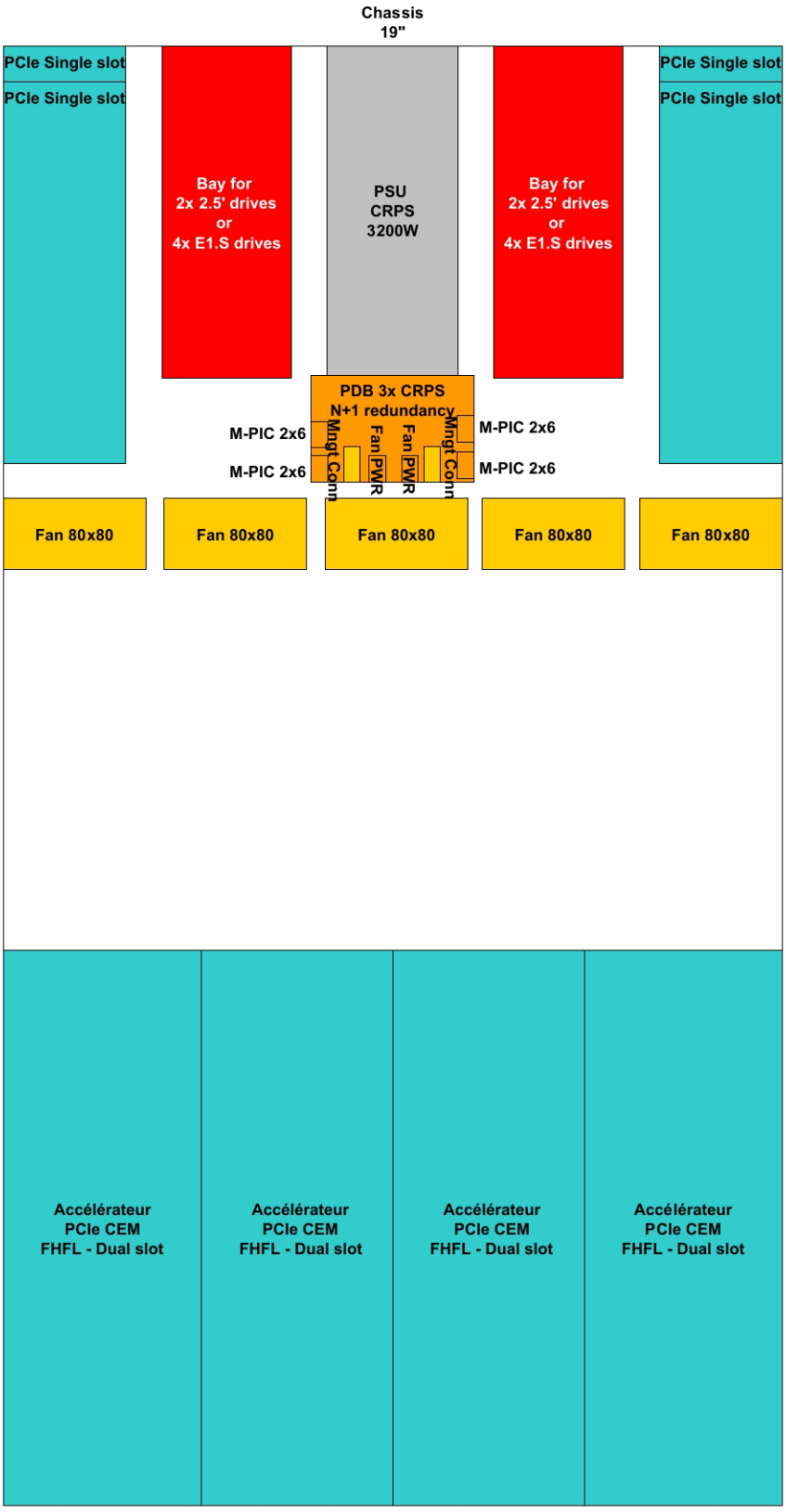

FIGURE 11 – M-FLW UPPER U OPTION 2

Figure 10 and 11 represent the second U of the DC-MHS server with M-FLW motherboard. We can see two different options.

The back of the server will integrate similar devices as we have in the first U of the server. We have different storage options (E1.S, 2.5" drives or M.2 drives with adapter) in a 5.25" bay or half 5.25" bay drive. If a half 5.25" bay drive is used, we can add a PCIe slot to connect different devices (network card, RAID card, HBA card, etc…).

The second CRPS PSU is located at the second U and connected to the same PDB as the CRPS PSU in the first U. The CRPS are connected to one PDB, which is placed vertically, while the CRPS are placed horizontally.

For higher performance, we will have two solutions:

- Solution based on OAM: for this solution we will be able to integrate two EPAC/EUPILOT HPM modules developed in the HIGHER project. Those modules have M-SDNO Class A335 form factor and can each integrate two OAMs. See chapter 3.4 for a more detailed description of the EPAC/EUPILOT HPM. The front of the server will be specifically designed for this option. Each module will be connected to the motherboard through cables on the MXIO connectors available on each of the SDNO Class A HPM. Those cables will be inserted in MCIO or MXIO connectors available on the motherboard.
- Solution based on standard PCIe CEM form factor: four PCIe 16x slots will be available to connect PCIe CEM form factor devices. Those devices can be FHFL (Full Height Full Length) and dual slot, making it possible to connect any type of device. The objective is to be able to integrate the latest graphic cards (GPU) available on the market, with specific care taken for their cooling since the TDP of the latest GPUs can be high (up to 600W). Being able to integrate 600W GPUs is a target but some more thermal study needs to be done to qualify compatibility with TDPs that high and cooling capacity for 600W in air cooling.

3.1.2 M-SDNO server

HPM compatibility

This second server will be compatible with the M-SDNO form factor from OCP for the HPM. The objective is to reuse as many building blocks from the M-FLW HPM server as possible to increase compatibility and interoperability between the two servers.

There are 5 different form factors in the M-SDNO specification. There are Class A, B, C, D and E. Class D HPMs are only compatible with 21" servers so it won't be possible to integrate them in the M-SDNO server since it is 19".

Class A, B and C are interoperable, which means that we can use those 3 different form factors without having to modify the chassis. The M-SDNO HPM server will only be compatible with Class A, B and C. Class E is intended for 19" but isn't compatible with Class C HPM. The main difference between Class C and Class E is that Class E integrates the power connectors so that we can plug PSUs directly into the HPM module while Class C provides some connectivity for OCP NIC 3.0, DC-SCM and storage and needs to have the power deported with PSUs connected to a PDB and the HPM being powered up via cables.

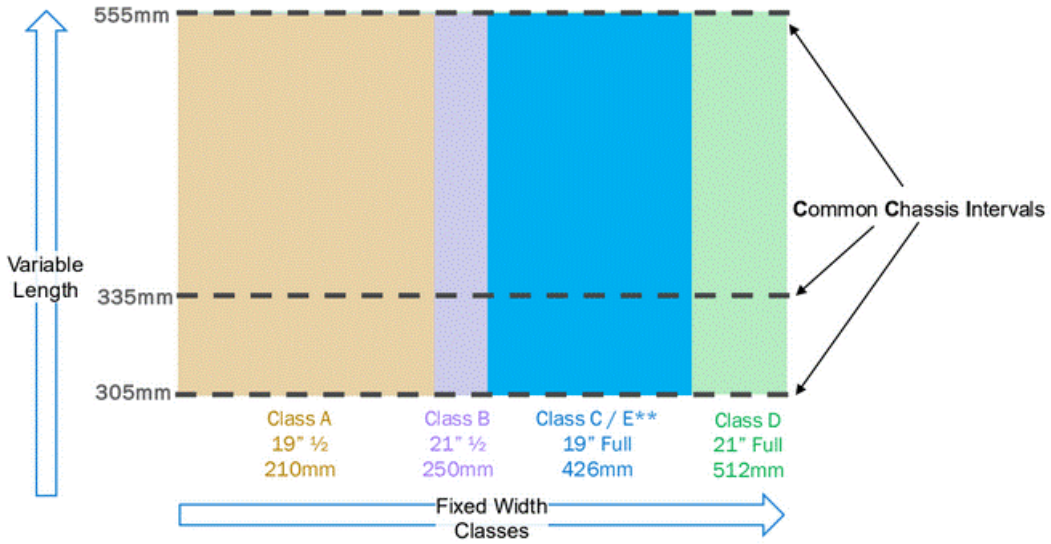

FIGURE 12 – M-SDNO CLASSES AND FORM FACTORS

As we can see in figure 12, the length is also variable. For our use case, the chassis will be compatible with a length of 335mm.

For the HIGHER project, the HPM for dual socket Rhea2 will be M-SDNO C335, while the single socket Rhea2 HPM and the EPAC/EUPILOT HPM will have the M-SDNO A335 form factor. Having those two form factors means that it will be possible to have either one M-SDNO Class C in the chassis or two M-SDNO Class A, without having to change anything at the chassis level. The fixing holes will be the same, the MXIO position will be similar (allowing even an interoperability for the cabling).

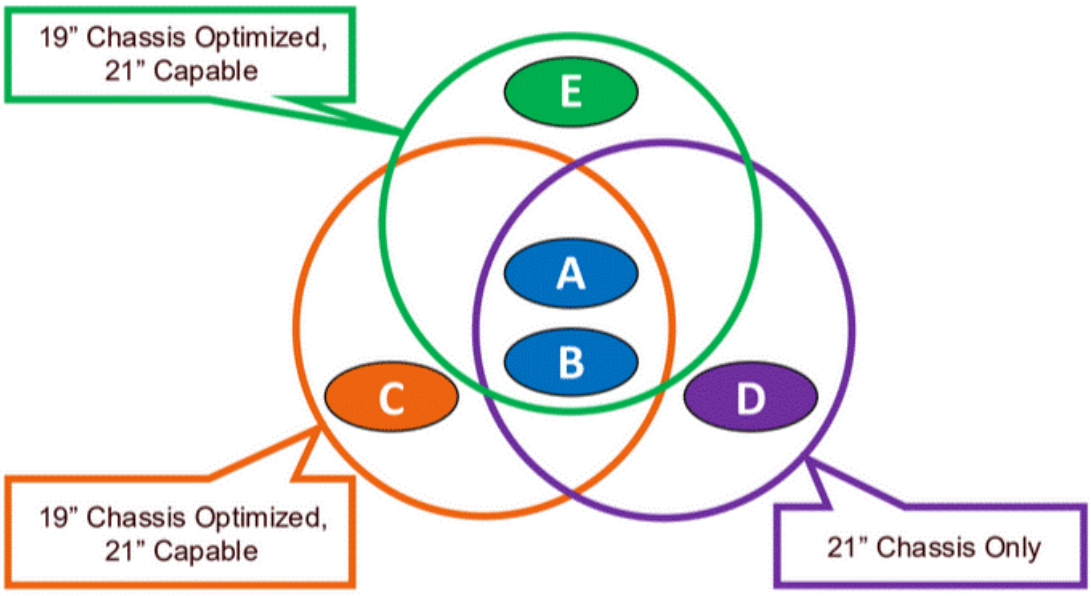

FIGURE 13 – M-SDNO INTEROPERABILITY

On the M-SDNO Class A HPM, we have two slots at the front, as we can see in figure 14. One slot is used for DC-SCM and one slot for one OCP NIC 3.0. The connectors on the Class A HPM will follow the M-PIC specification from OCP.

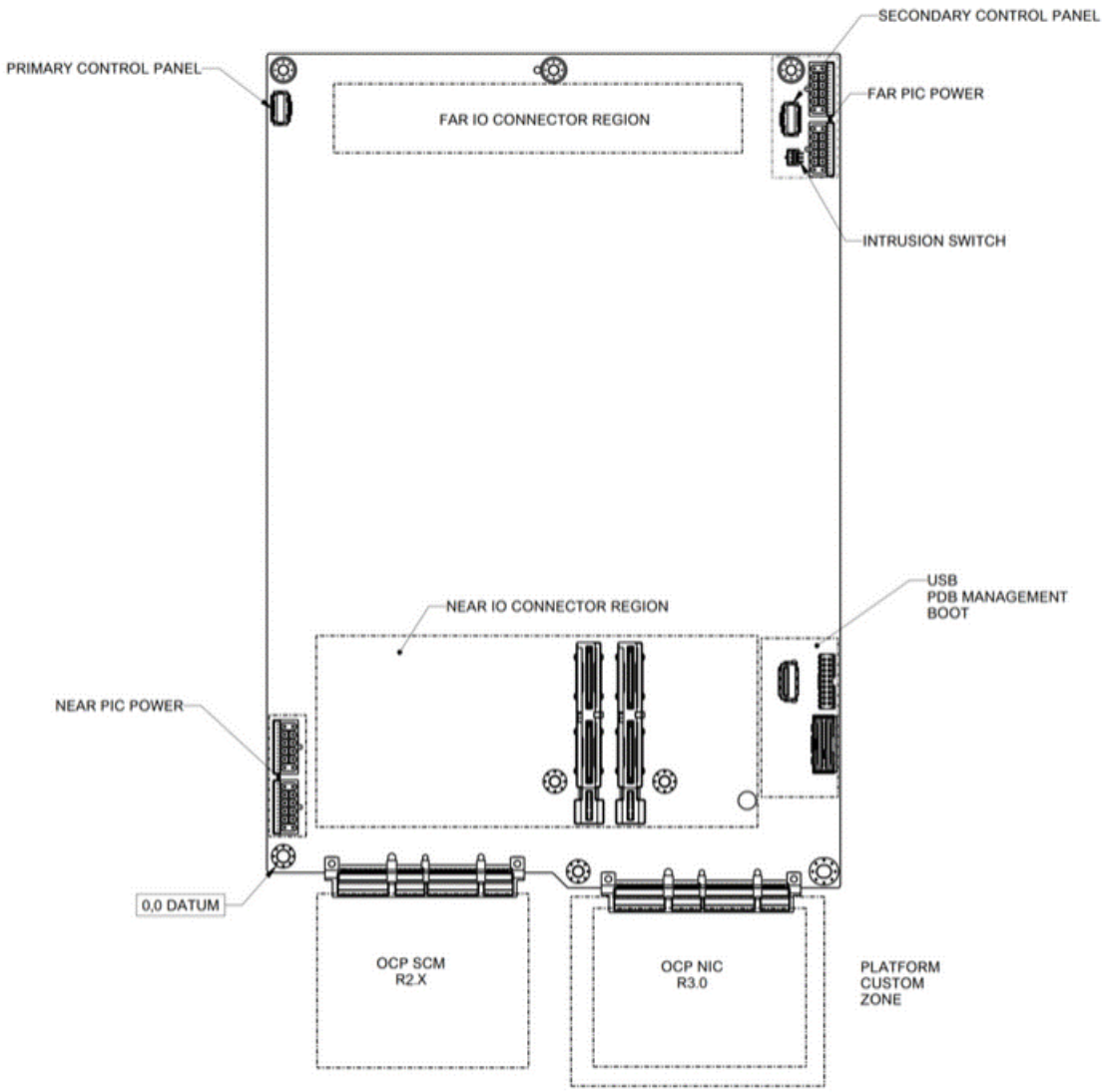

FIGURE 14 – M-SDNO CLASS A HPM OVERVIEW

On the M-SDNO Class C, we still have one slot for the DC-SCM module and one slot for one OCP NIC 3.0, and we also have a platform custom zone. This zone can integrate different devices depending on the board manufacturer. It can be either a second slot for one OCP NIC 3.0 or slots for storage (like E1.S drive with M-FLW form factor). The other connectors on the Class C HPM also follow the M-PIC specification so we will have the same type of connectors on the Class A and Class C HPM.

If a second OCP NIC 3.0 slot is required for the M-SDNO Class C HPM, it will be at the same position as the OCP NIC 3.0 would be if we were using two M-SDNO Class A HPMs, as we can see in the figure 15 below:

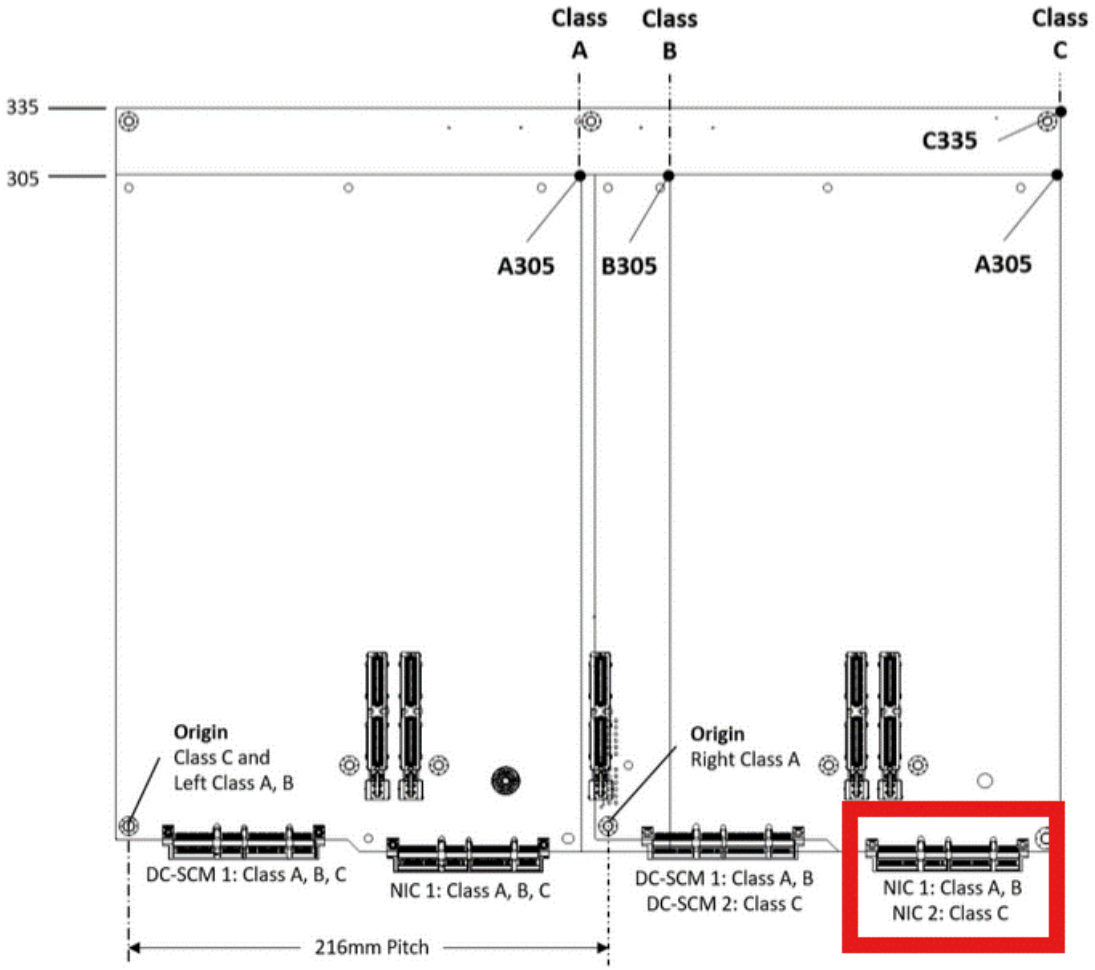

FIGURE 15 – M-SDNO CLASS C OVERVIEW

Power

The power supply part will be similar to the M-FLW HPM server. The M-SDNO server is going to be compatible with Class A, B and C. None of those classes integrate the power connector directly into the HPM like it is the case for M-FLW or M-SDNO Class E. A Power Distribution Board will be required to be able to power up every component in the server. The same PDB designed for M-FLW chassis can be reused in the M-SDNO chassis. Two CRPS, each one up to 3200W, will bring the N+1 redundancy to the server, regardless of whether we are using one M-SDNO Class C or two M-SDNO Class A. Power supplies will be inserted horizontally while the PDB will be in a vertical position.

Internal architecture and connectivity

M-SDNO Class C HPM

When used with M-SDNO Class C HPMs, the server's main use case will be for compute, with different types of storage and some PCIe slots to connect different types of devices (small GPU, network card, RAID card, etc…).

For the storage, the same bays as in the M-FLW server can be used so it will be possible to have 2.5" drives and E1.S drives in half or full 5.25" bay. It will be possible to have M.2 drives directly connected to the HPM in M-SDNO form factor when used as host and not as accelerator module.

Connectors at the front will be available for interconnection between different DC-MHS systems.

The cooling within the systems will be the same for both configurations, with one M-SDNO Class C or two M-SDNO Class A. It will be the same solution as for the M-FLW server with several fans of 80x80mm, managed by the system. They will be powered up by the PDB.

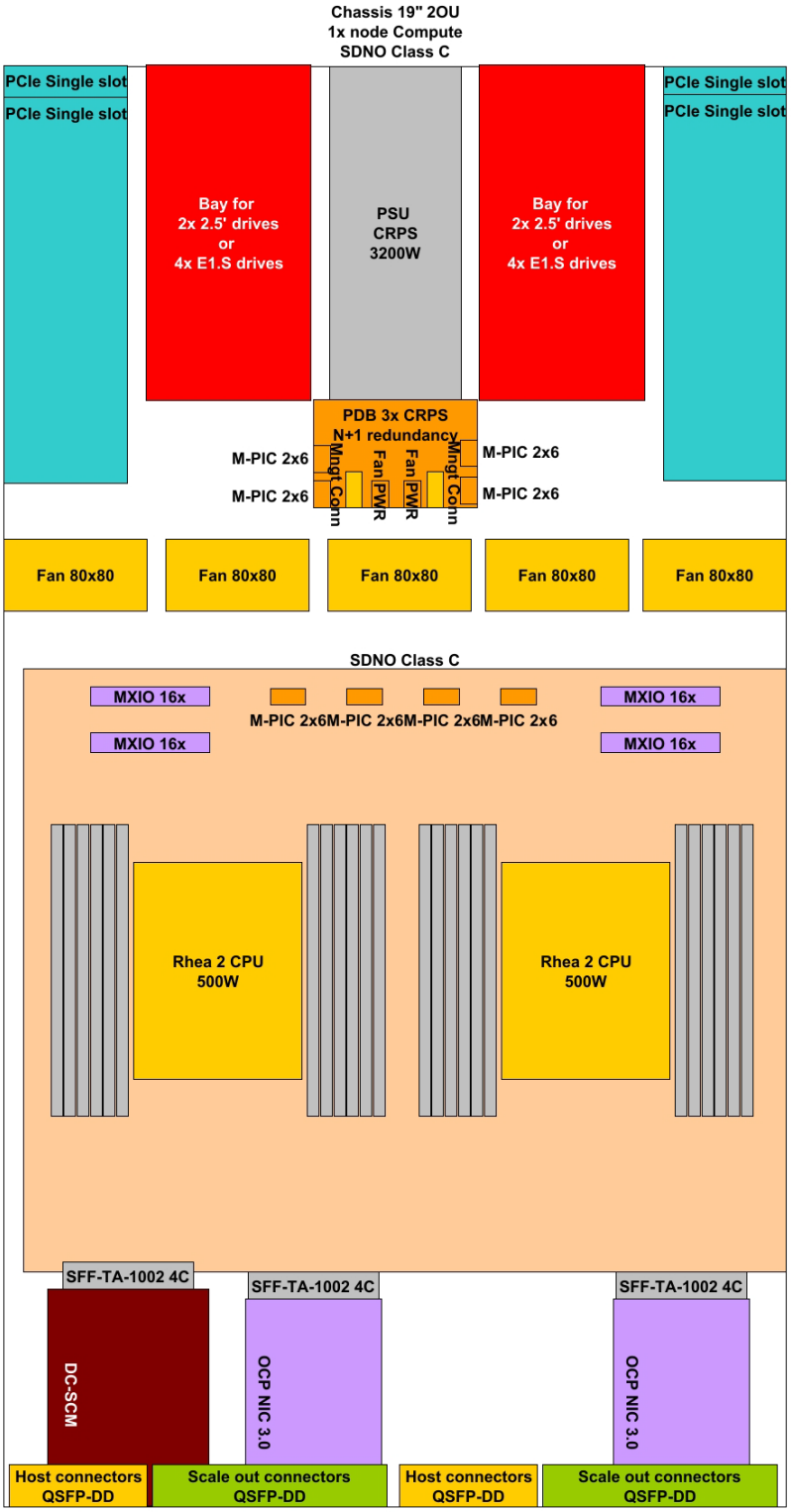

FIGURE 16 – M-SDNO CLASS C, LOWER U

In figure 16, we can see the first U of the DC-MHS server with M-SDNO Class C motherboard. At the back, we have the different storage options and some PCIe slots for network cards, RAID cards, HBA cards, etc… As for the M-FLW DC-MHS server, PSUs will be on top of each other in the center of the server so that we can easily move to a 21" form factor server in the future by adding rails to the server and a cable that we connect to the busbar to follow [ORV3] specification. Both PSUs are connected to the PDB, which powers up the different components of the server (motherboard, PCIe card, fans, …). The M-SDNO Class C HPM will integrate two Rhea2 processors (See 3.2 Dual socket Rhea2-based HPM). Each processor can have a TDP of up to 500W. A 2U form factor heatsink will be required to dissipate the heat coming from the CPU. The PCIe 16x connectors on the motherboard will be connected to the different components in the server, or outside of the server, with cables. One DC-SCM and two OCP NIC 3.0, one per CPU, will be accessible at the front of the server. Scale out connectors and host connectors (most probably QSFP-DD connectors, but this could change during the development of the project) will be used for the interconnection between DC-MHS servers.

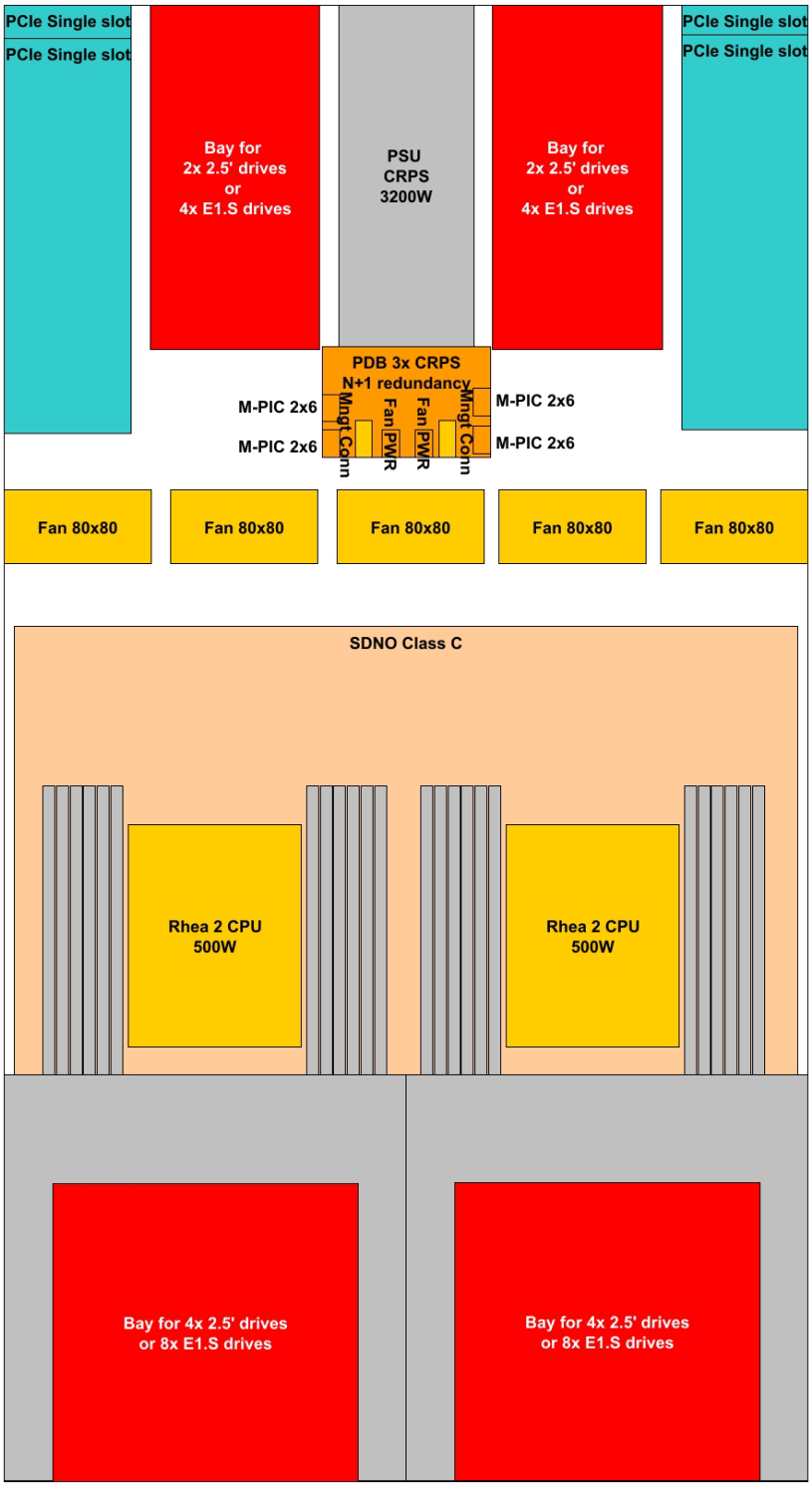

FIGURE 17 – M-SDNO CLASS C, UPPER U

At the second U level, that we can see in figure 17, we have some more storage accessible at the front of the server and like in the first U, at the back we can have different options for storage and/or PCIe slots. The second PSU is at the second U level. Furthermore, we will have the CPU's heatsink and fans in this level as well since they will have a 2U form factor.

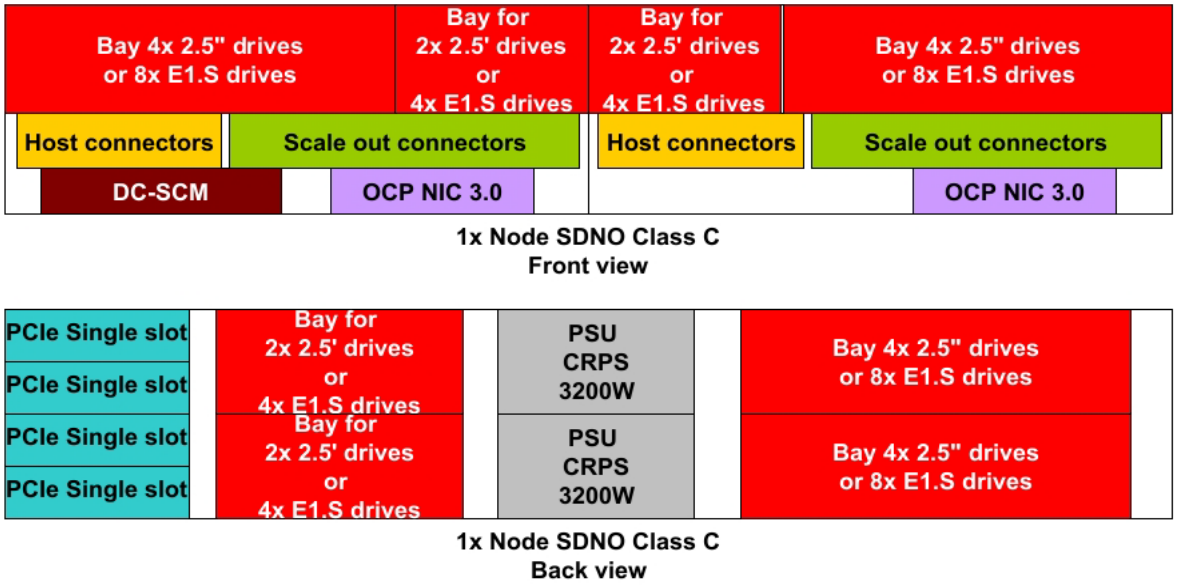

FIGURE 18 – FRONT AND BACK VIEW OF M-SDNO CLASS C SERVER

In figure 18, we have the front and back view of the server with the different components. Several configurations in terms of storage and PCIe slots availability are possible, depending on the requirements and use cases. The scale out connectors and host connectors are used with interconnections between the different DC-MHS systems that are going to be developed for the HIGHER project. It would be complicated to replace two PCIe single slots by a single PCIe dual slot since dual slot PCIe cards usually come in FHFL (Full Length Full Height) format and it would be really tight to insert them between the chassis and the PSUs in the centre of the server, as we have to also take into consideration the cabling and the size of the riser on which the PCIe cards are inserted.

M-SDNO Class A HPM

The server will be compatible with different types of M-SDNO Class A HPMs:

- M-SDNO Class A for x86 CPUs, from different manufacturers (ASrockRack, Compal, MSI, etc…)
- M-SDNO Class A with single socket Rhea2 CPU that will be developed during the project (see 3.2 Dual socket Rhea2-based HPM)
- M-SDNO Class A with two EPAC/EUPILOT modules, with the modules having the possibility to act as host or not (see 3.4EPAC/EUPILOT Processor Module).

It will be possible to mix those different types of M-SDNO Class A HPMs within the same chassis.

The same chassis as for the M-SDNO Class C HPM will be used. Power and storage options will be the same. The two CRPS will provide all the power necessary for the different HPMs. Depending on the power requirement for the HPM that will be used, it will be possible to have N+1 redundancy, since the CRPS is able to go up to 3200W. The storage options will be the same as before, with the possibility to have M.2 modules, 2.5” drives or E1.S drives in the server (either directly on the HPM or in 5.25” bay drives and half 5.25” bay drives). PCIe slots will be available, depending on the use case and the PCIe lanes available coming from the HPMs. 80x80mm fans will be used for the cooling solution of the different types of M-SDNO Class A HPMs that are going to be used.

The main difference will be at the M-SDNO Class A HPM level. Depending on the type that is used, we can cover different use cases. For each configuration, the second U will be the same. We will have

the same options as with M-SDNO Class C in terms of storage and PCIe slots. The back of the server will be the same as for M-SDNO Class C version. For the front, there will be one more slot available for the DC-SCM board of the second M-SDNO Class A HPM.

It will be possible to use 2U form factor heatsinks on the CPU and EPAC/EUPILOT module present on the M-SDNO Class A HPM. The choice between 1U or 2U heatsink will depend on the TDP of the CPU and EPAC/EUPILOT module and cooling requirements. The goal is to be able to dissipate as much heat as possible passively to be more efficient for power efficiency in the server. With a bigger heatsink, we could reduce fan speed and use less power in the server.

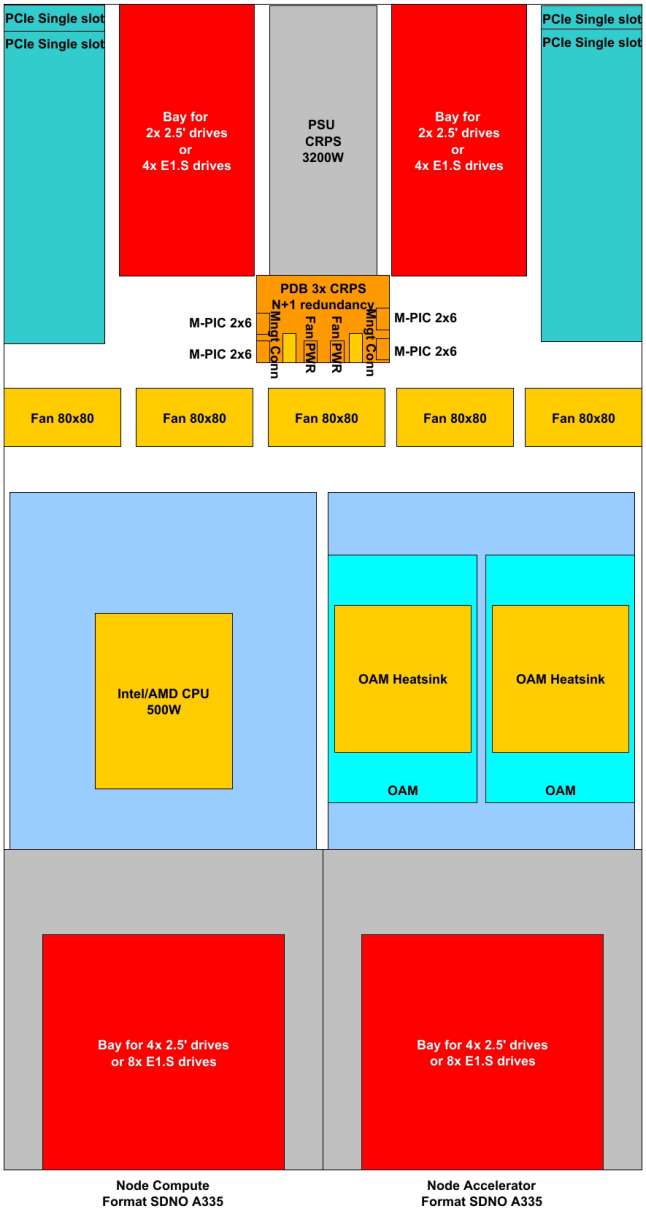

FIGURE 19 – M-SDNO CLASS A HPM, UPPER U

In figure 19, we have the second U of the M-SDNO Class A server. The only difference is for the heatsinks that will be used, otherwise we have the same configuration as what is planned for the M-SDNO Class C version.

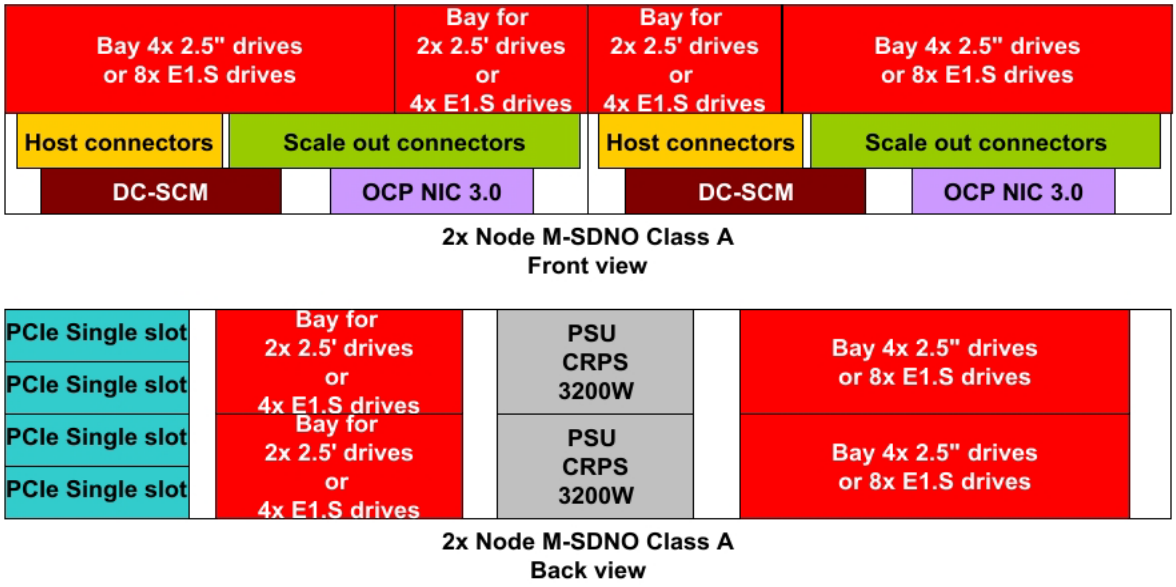

FIGURE 20 – M-SDNO CLASS A - FRONT AND BACK VIEW

The back as well as the front of the server, that are depicted in figure 20, will almost be exactly the same as for M-SDNO Class C HPM except for the second DC-SCM slot that we have for the second M-SDNO Class A HPM (only one DC-SCM slot is required for M-SDNO Class C).

M-SDNO Class A HPMs with standard x86 usually come with several MCIO 8x connectors on board (for PCIe connectivity), usually placed at the far side and the near side. Storage is available at the motherboard level with up to 2x M.2 slots.

FIGURE 21 – GNRAPD12DNO FROM ASROCKRACK

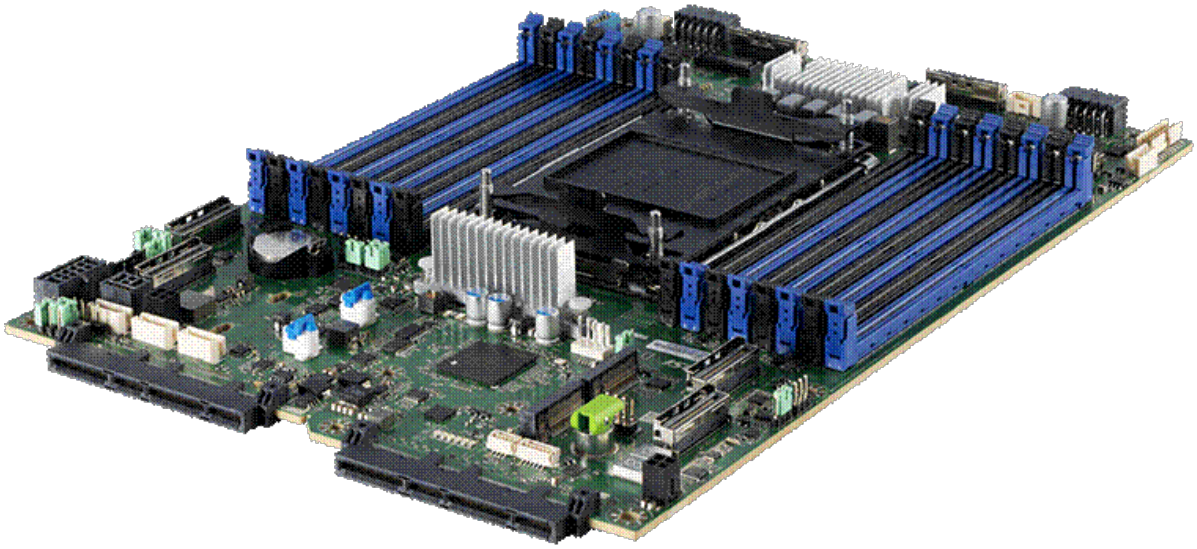

FIGURE 22 – D3061 FROM MSI

In figure 21 and 22, we can see two examples of boards that are available in the market. Those boards have the M-DNO Type 2 form factor, which is really similar to M-SDNO Class A. The main difference is that for M-DNO Type 2, the position of several components on board are fixed while for M-SDNO Class A, we don't have a fixed position for the connectors but a zone in which they should be implemented. The dimensions and fixing holes are the same between the two form factors. On both motherboards, we can see the different MCIO connectors for the connectivity to other devices within the server and the two slots at the front for one DC-SCM board and one OCP NIC 3.0.

It will be possible to have two such Class A motherboards in a chassis. The target is to connect one to the EPAC/EUPILOT processor module like we can see in figure 23 below:

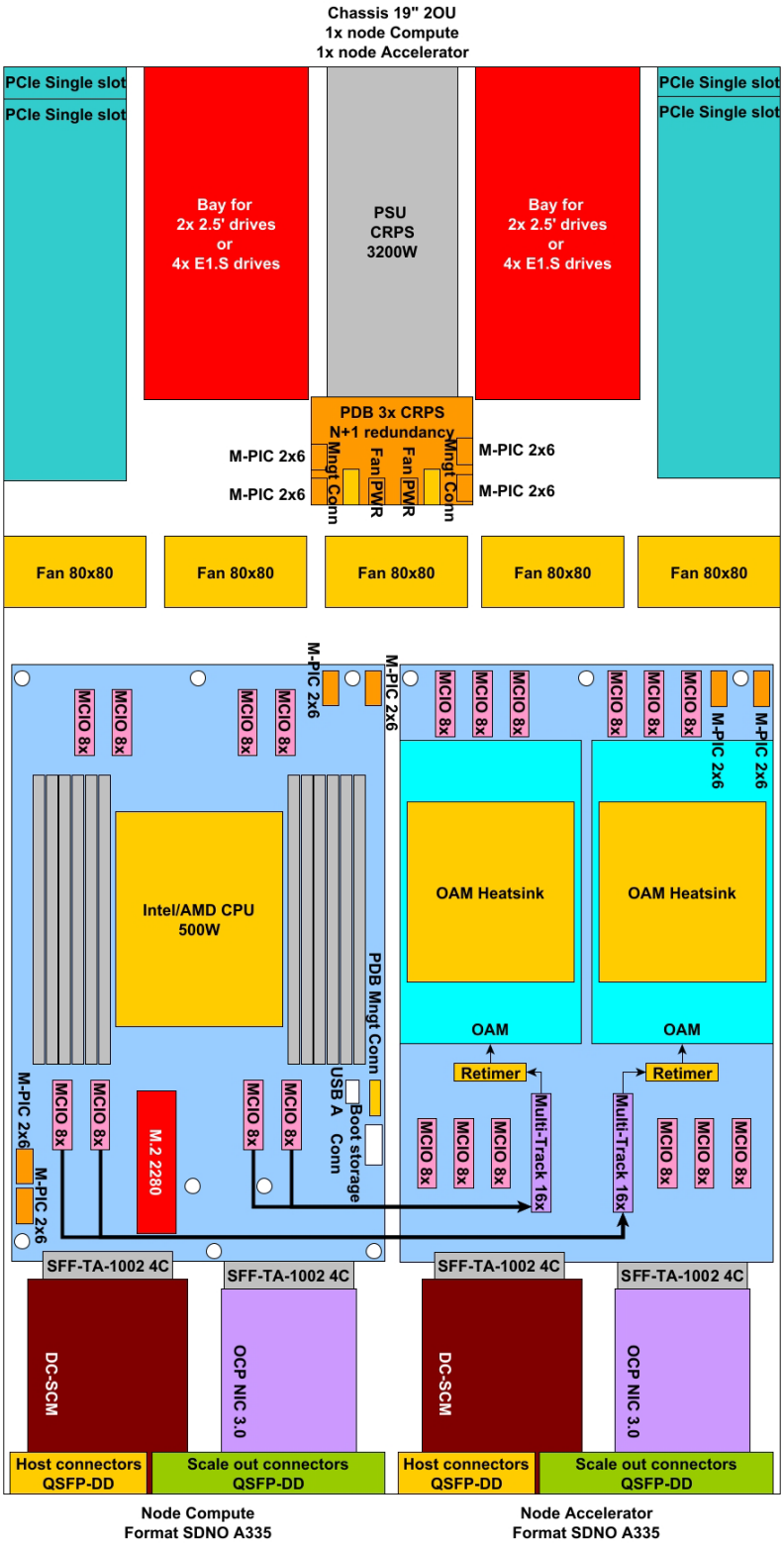

FIGURE 23 – SDNO CLASS A NODE COMPUTE + EPAC/EUPILOT

Both HPMs will be connected with cables, going from MCIOs available on the motherboard with the CPU to the Multitrack 16x on the EPAC/EUPILOT processor module. M-PIC connectors on the boards

will allow powering up of both HPMs. Cables between PDB and HPMs will be used for power distribution.

The version with the single socket Rhea2 based HPM will be the same as for the x86 based HPM. If MCIO 16x are available on the Rhea2 based HPM, it will be possible to have the connection with the EPAC/EUPILOT processor module with MCIO 16x to MXIO 16x cables. A cable with two MCIO 8x to one MXIO 16x can also be used if required.

In figure 24 below, we have the version with two EPAC/EUPILOT processor modules. This option must be connected to another HPM server on which we will have the host that will be connected to the accelerator modules. This host can be either a server with M-SDNO Class C or a server with one of the M-SDNO Class A acting as host. In the latter case, we may use either an x86-based HPM,a Rhea2-based HPM or a EPAC/EUPILOT processor module with one of the OAMs acting as a host. The interconnection of the servers is described in chapter 3.1.3.

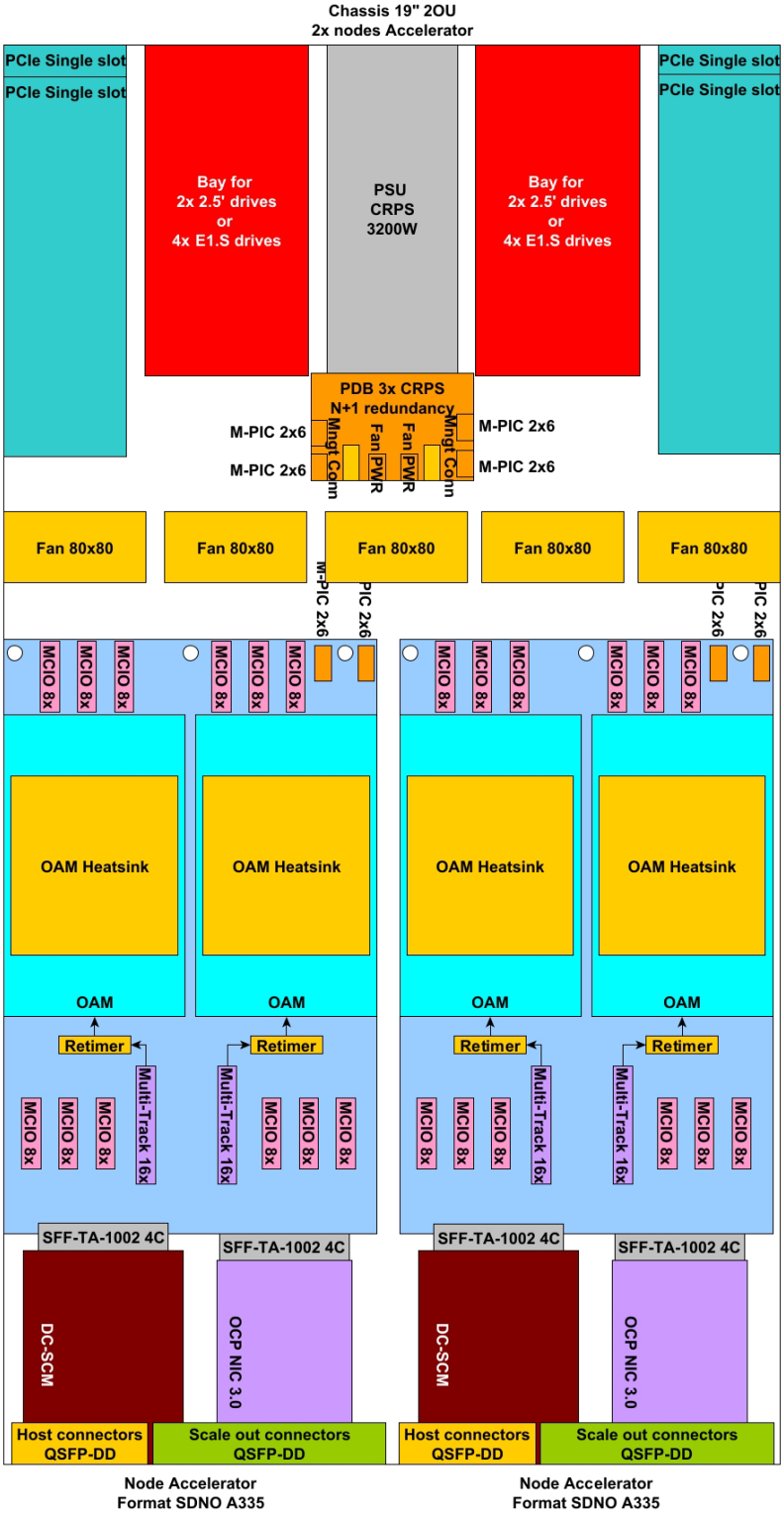

FIGURE 24 – SDNO CLASS A - TWO EPAC/EUPILOT HPMS

The last configuration, shown in figure 25, will be also with two EPAC/EUPILOT processor modules, but one of the OAMs will be acting as a host. The two EPAC/EUPILOT processor modules will be connected to each other with different cables. Those cables will be MCIO to MCIO and MXIO to MXIO.

The OAM acting as a host will be the only difference to the previous configuration (two EPAC/EUPILOT processor modules with no host). The options for storage, PCIe slots and power will be the same. The front of the server will also be the same.

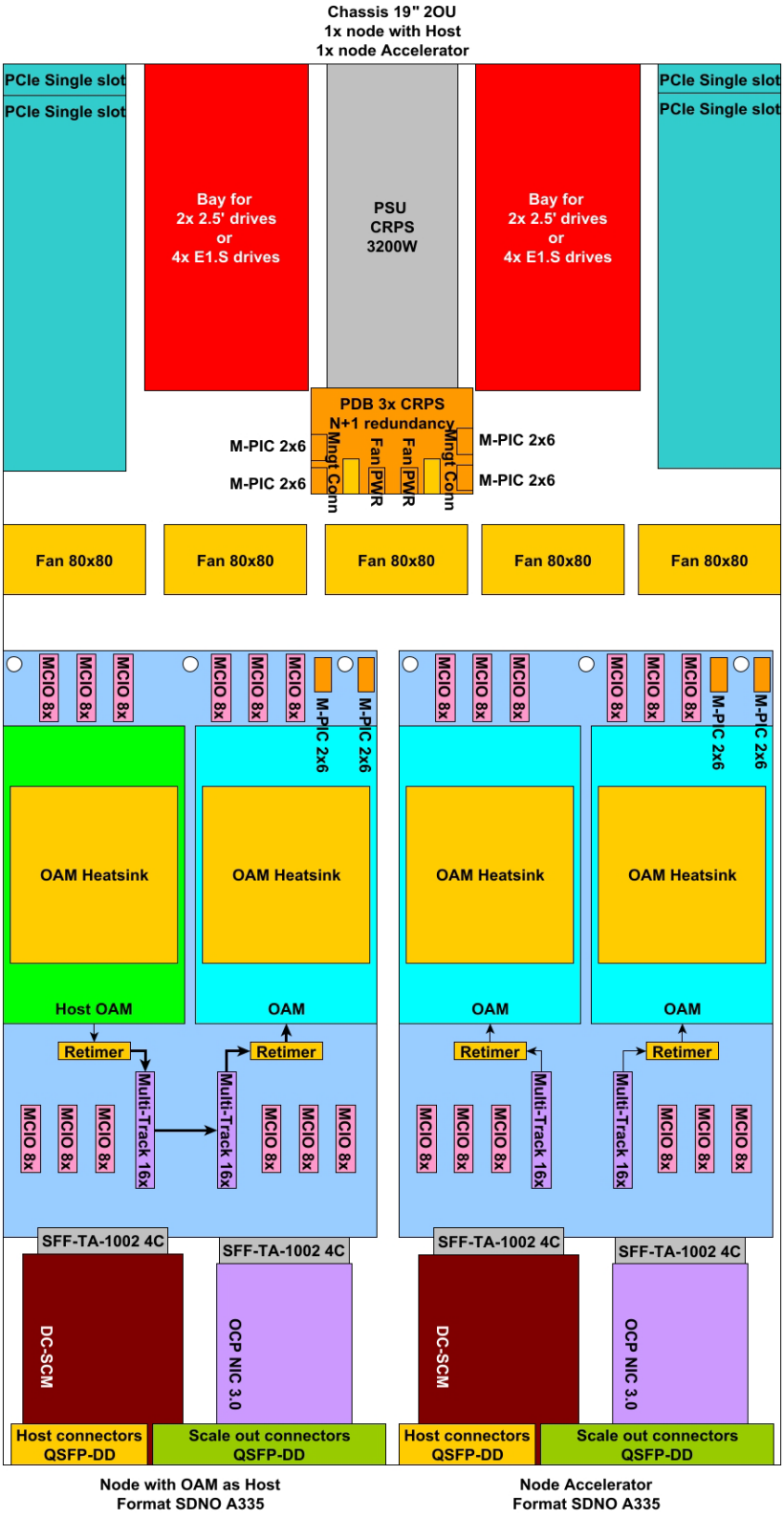

FIGURE 25 – M-SDNO CLASS A WITH ONE OAM AS HOST

3.1.3 Interconnection between DC-MHS servers

It will be possible to interconnect several DC-MHS servers.

Connectors at the front of the servers will be used for the connectivity. Those connectors will have two different roles.

- Host connectors: PCIe lanes coming from the host will go through the host connectors. It should be possible to have at least 32x PCIe lanes to connect to the 2x MXIO 16x that we can have on the EPAC/EUPILOT processor module.
- Scale-out connectors: those connectors' goal is to be able to interconnect OAMs that are in different HPM servers. The target is to be able to have up to six OAMs connected to each other. One objective for the scale out connectors is to be able to handle up to 48x PCIe lanes.

The connectors that have been selected for the moment are QSFP-DD connectors, which can handle up to 8x PCIe lanes per connector. Depending on what will be available in the market, it is possible that other connectors are used in the future. Compatibility with the HPM will be assured and validated if other connectors are preferred.

QSFP-DD connectors usually come in bays of 2x, 4x or 6x connectors. The bays will be connected to a specific Printed Circuit Board (PCB) which will operate the connection between the HPM and the QSFP-DD connectors (via internal cables).

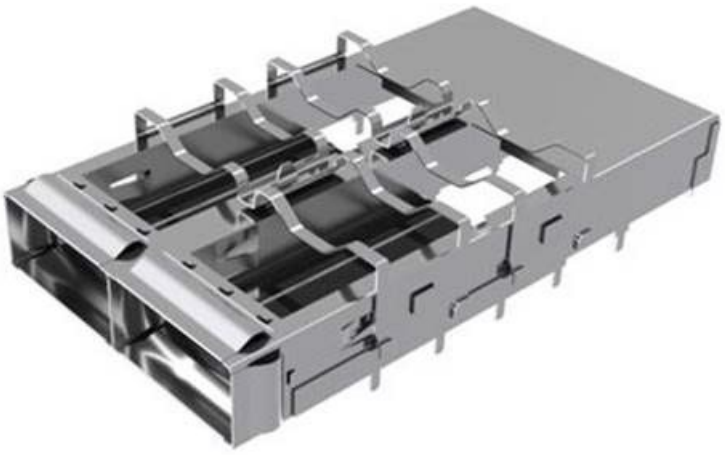

FIGURE 26 – QSFP-DD BAY FOR TWO CONNECTORS

FIGURE 27 – QSFP-DD BAY FOR FOUR CONNECTORS

On figure 26 and 27, we can see examples of bays for 2x and 4x connectors. Bays for 6x connectors are available on the market, but less common than the bays for 2x and 4x.

Cables will be used to connect the QSFP-DD connectors from different DC-MHS servers, depending on the use case and requirements.

The connectors at the level of the DC-MHS servers and cables allow a lot of modularity between the systems. With this solution it will be possible to have different topology and connectivity between host and the EPAC/EUPILOT processor module.

3.2 Dual socket Rhea2-based HPM

As a reminder, Rhea2 features a dual chiplet solution connected through advanced Universal Chiplet Interconnect Express (UCIe) die-to-die links with:

- 128 Arm Neoverse V3 cores with SVE
- Arm9.x-A series ISA
- Advanced Arm CMN-S3 Mesh fabric (NoC)
- PCIe Gen 6 (96 lanes total) and CXL 3.0 support
- 12 DDR5 channels to RDIMM for optimized data throughput

3.2.1 Requirements

According to [D2.1], the dual-socket Rhea2-based HPM ([Rhea2-HPM]) is SDNO Class C335 compliant.

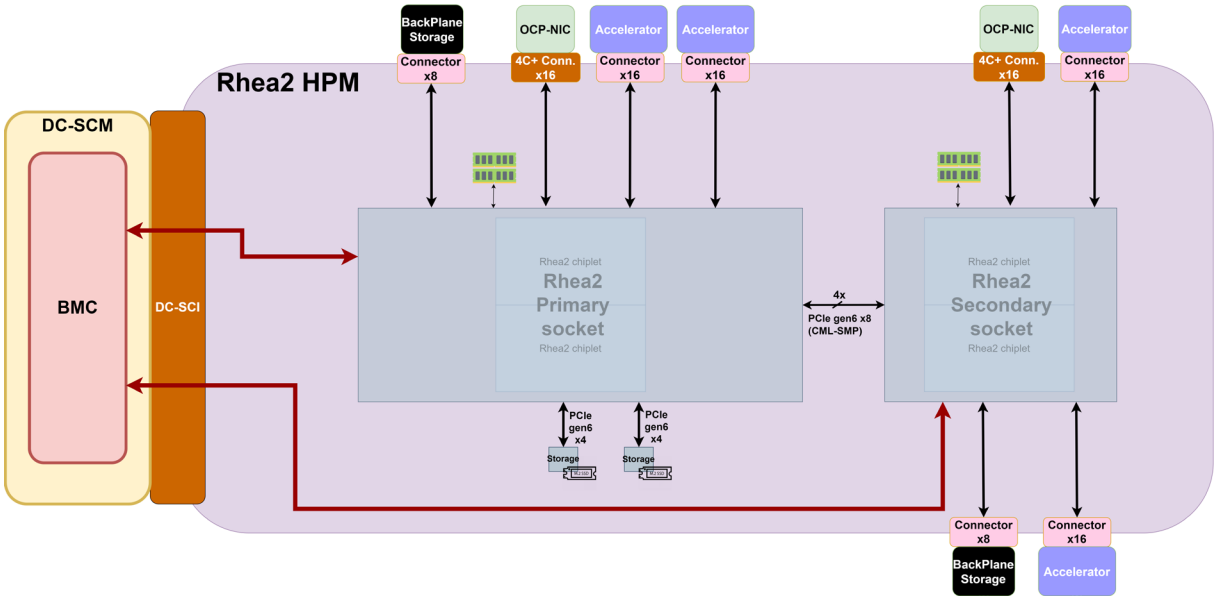

FIGURE 28 – DUAL SOCKET RHEA2-BASED HPM - HIGH-LEVEL BLOCK DIAGRAM

3.2.2 Main characteristics

The Main characteristics of the [Rhea2-HPM] are given in the table below:

DDR5 memory	Total of 24 RDIMM connectors 12 RDIMM connectors per socket Up to 128GB@6400MT/s per RDIMM
PCIe connectors	• 4x Connectors with PCIe x16 Gen6 RC PCIe links for accelerators • 2x Connectors with PCIe x8 Gen6 RC PCIe links for storage • 2x Connectors with PCIe x16 Gen6 RC PCIe links for OCP NICs • 2x Connectors with PCIe x4 Gen4 RC PCIe links for on board M2 SSD
DC-SCM connector	• 1x LTPI • 2x UART • 3x I2C • 2x I3C • 2x PCIe x1 Gen4 RC

	• 1x USB2.0 • 1x USB3.0 • 2x QSPI • 1x eSPI
Power connectors	4x Connectors (2x6 + 12sb vertical header) 12V/12A per pin
Mechanical	SDNO class C (426 x 324 mm²)

TABLE 1: DUAL SOCKET RHEA2 MAIN CHARACTERISTICS

3.2.3 Preliminary layout

The Figure 29 – Dual Socket RHEA2-Based HPMFigure 29 below is a preliminary placement of [Rhea2-HPM]:

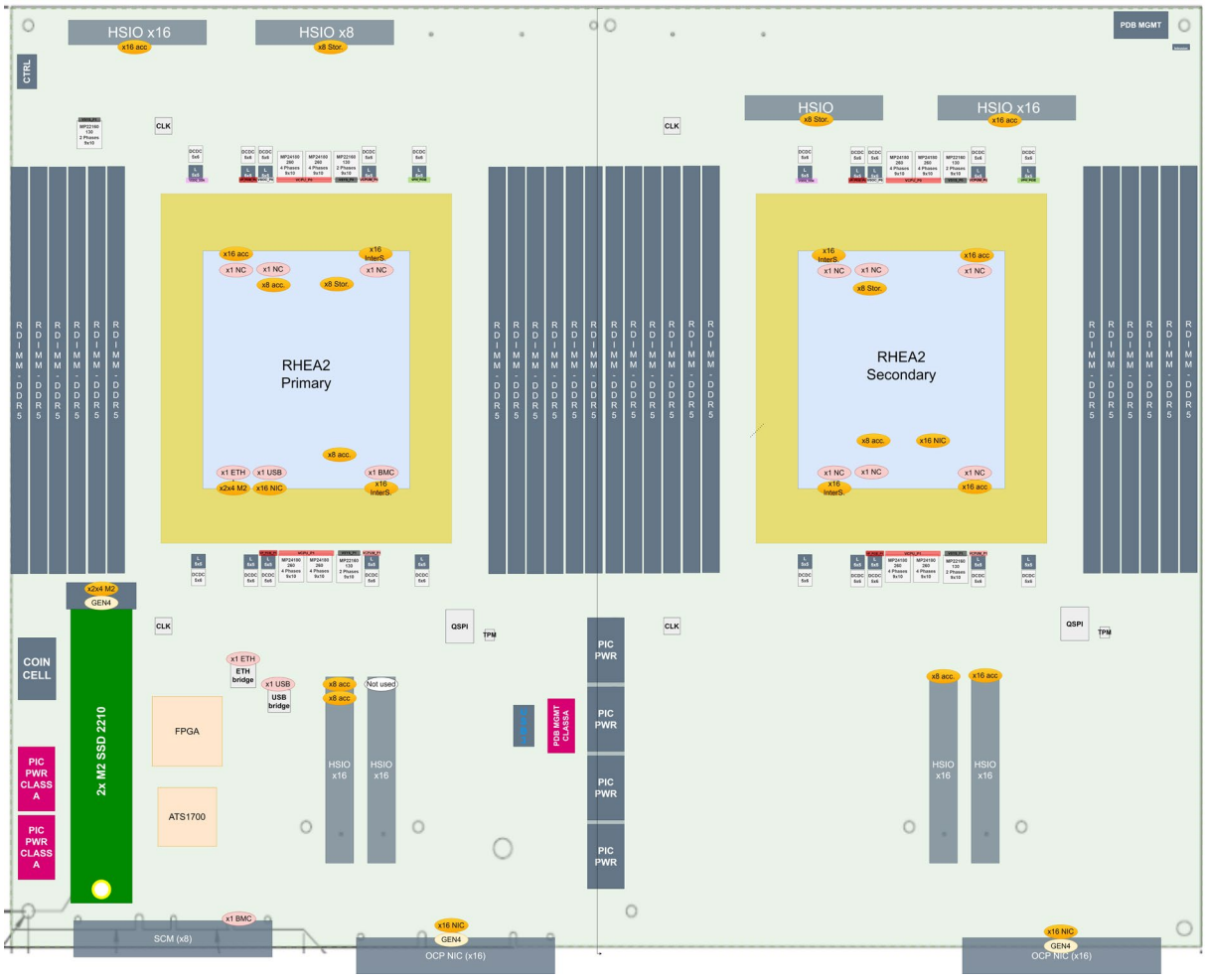

FIGURE 29 – DUAL SOCKET RHEA2-BASED HPM - PLACEMENT

3.2.4 Main interfaces

PCIe

All PCIe x16 interfaces are available on an HSIO connector, as defined in [M-SDNO] chapter 10.9.

All PCIe x16 interfaces support root complex and CML-SMP/CXL modes.

All PCIe x16 interfaces support 2x8, 4x4 n-Furcation.

Figure 30 is a description of the board's PCIe tree:

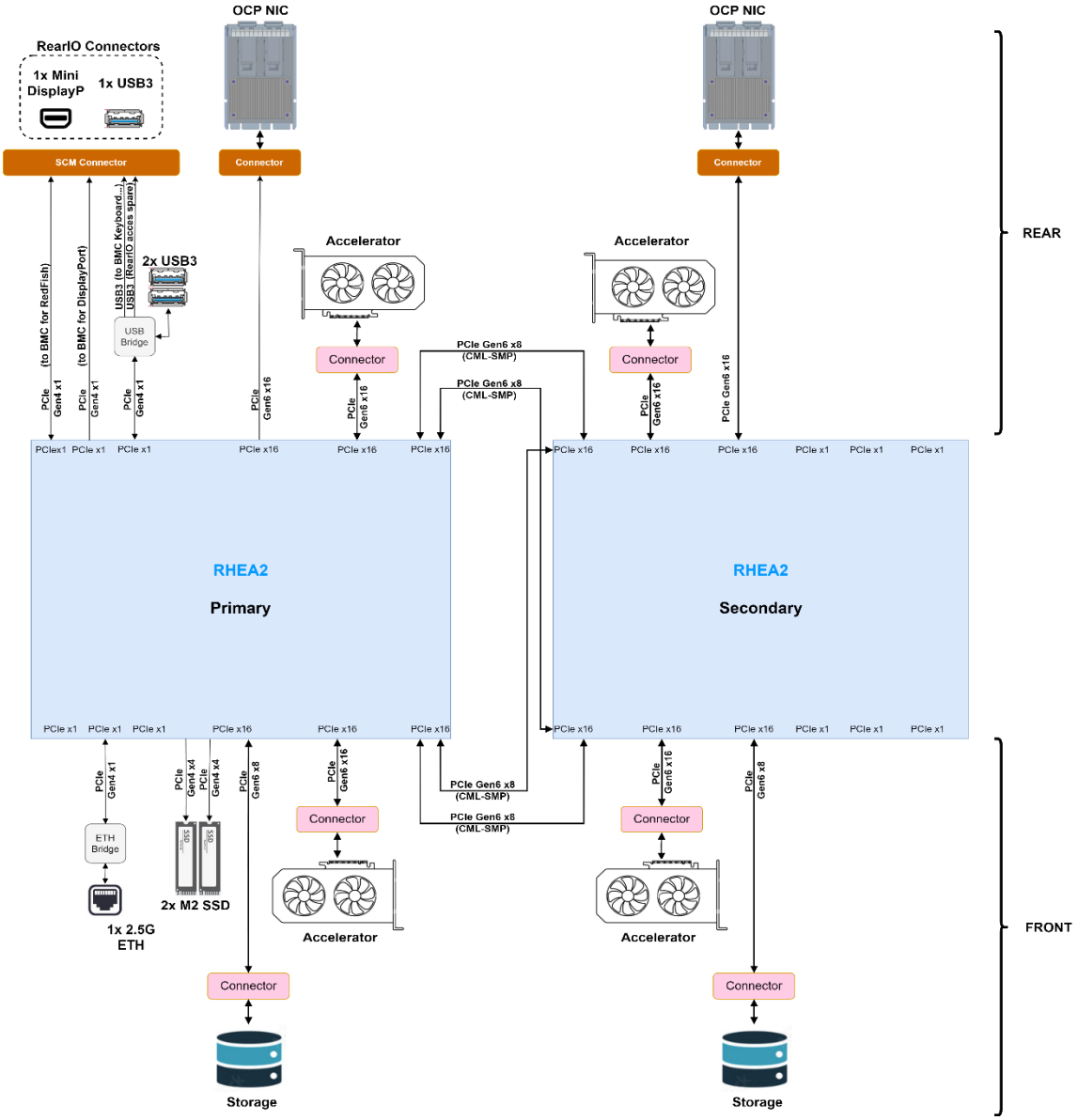

FIGURE 30 – DUAL SOCKET RHEA2-BASED HPM - PCIE TREE

Below is a summary of PCIe x16 functions:

– 4x PCIe x16 Gen6 connectors for accelerators (2 connected to primary socket and 2 connected to secondary socket)

- 2x PCIe x8 Gen6 for storage back planes (1 connected to primary socket and 1 connected to secondary socket)
- 4x PCIe x8 Gen6 CML-SMP for all to all inter die communication (2 connected to primary socket and 2 connected to secondary socket)
- 2x PCIe x16 Gen6 connectors for OCP-NIC
- 2x PCIe x4 Gen4 for M2 SSD connected to the primary socket (both are connected to the primary socket)
- 1x PCIe x1 Gen4 for USB3.0 bridge (UPD720201K8-701-BAC-A from Renesas) (connected to primary socket)
- 1x PCIe x1 Gen4 for ETHERNET bridge (I225 from Intel) (connected to primary socket)
- 1x PCIe x1 Gen4 for display port through DC-SCM's BMC (connected to primary socket)
- 1x PCIe x1 Gen4 for Redfish protocol (connected to primary socket)

Retimers from either Asteralabs or Broadcom will be used on PCIe x16 Gen6. They will be implemented on the backplanes and powered directly from the PDB (Power Distribution Board).

Side band signals will be available on the same connector as the high-speed signals.

Figure 31 below gives a more detailed view of how the PCIe x6 connectors and HPM-internal links will be involved to create the all to all connections between the 4 chiplets:

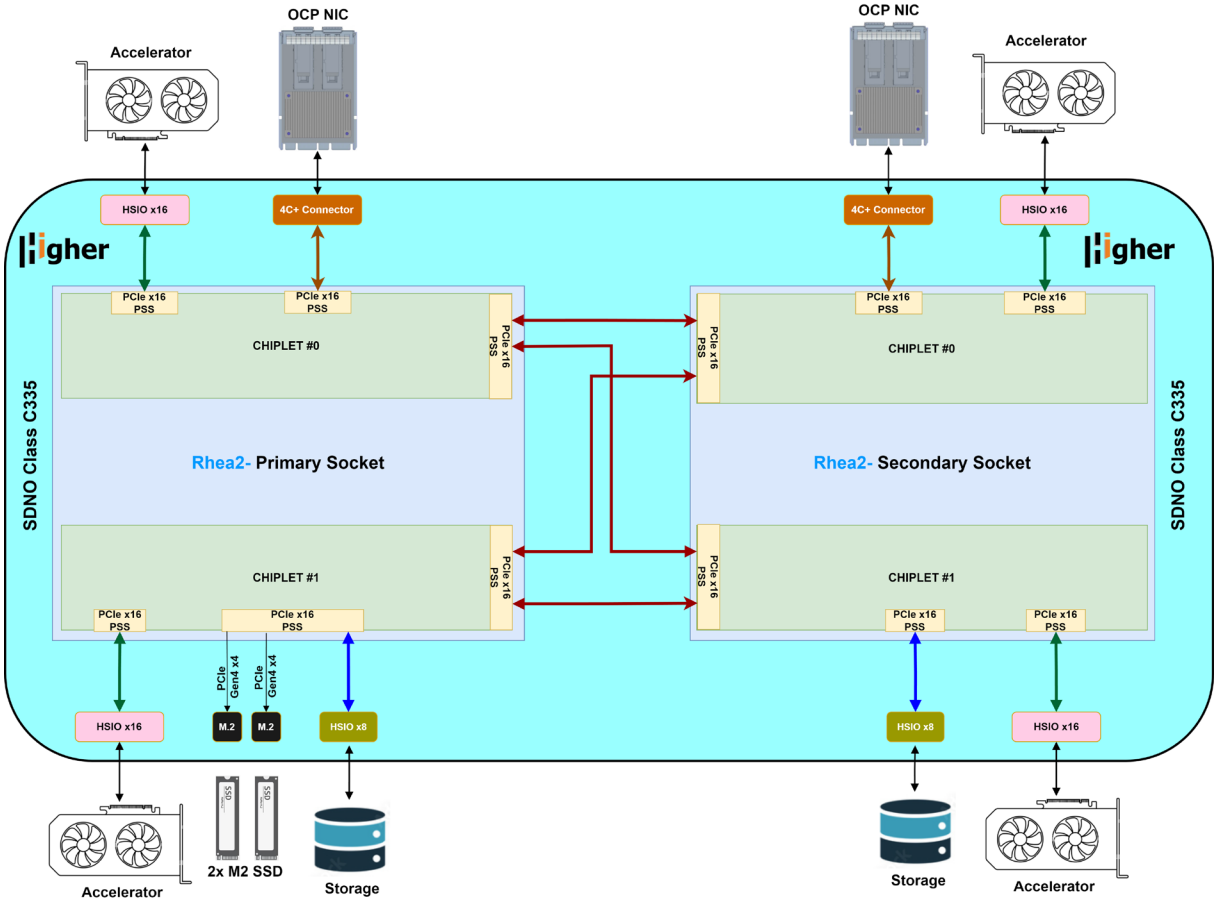

FIGURE 31 – DUAL SOCKET RHEA2 BASED HPM - ALL TO ALL BETWEEN CHIPLETS

DDR

Figure 32 and Figure 33 describe the DDR5 connections on RHEA2 primary and secondary sockets.

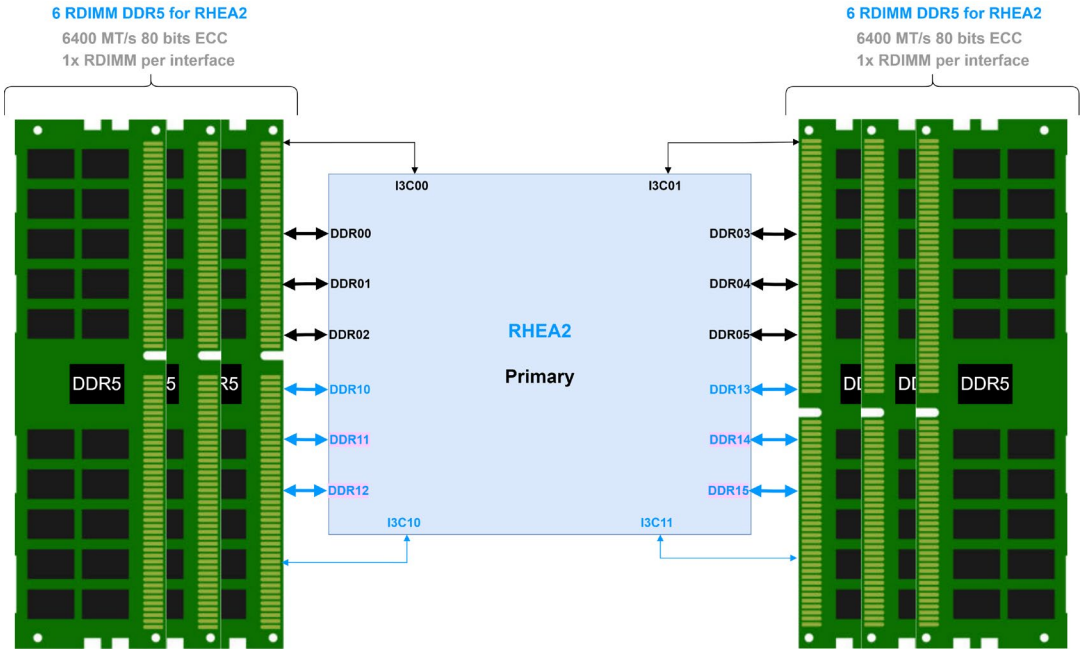

FIGURE 32 – DDR INTERFACES FOR PRIMARY RHEA2

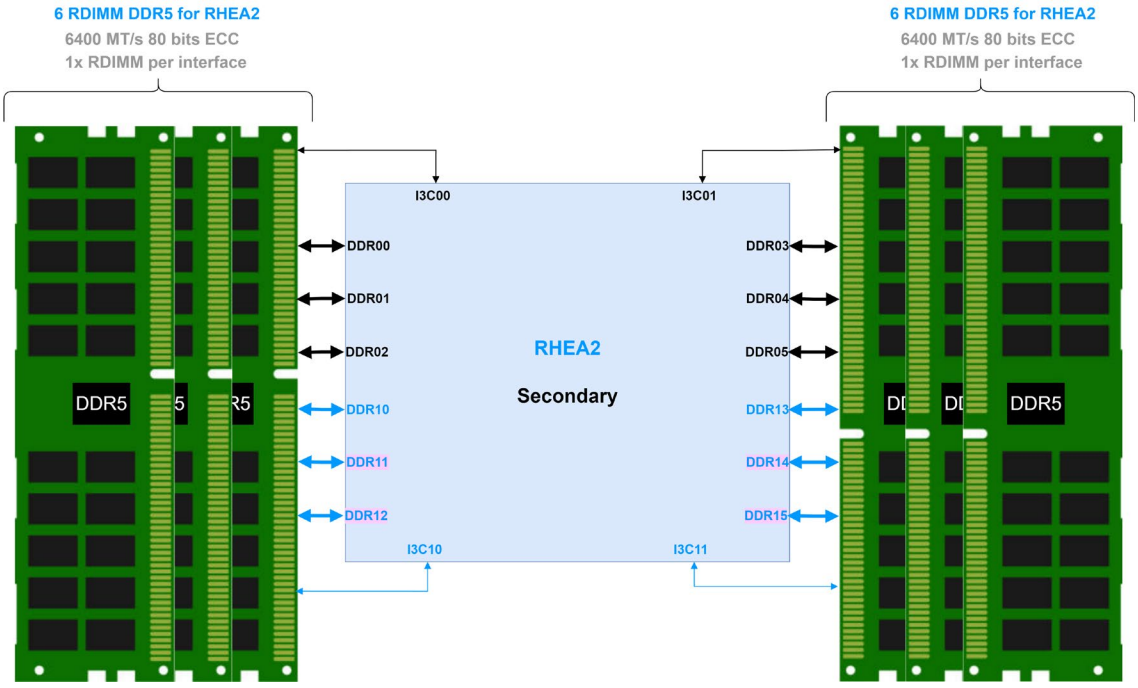

FIGURE 33 – DDR INTERFACES FOR SECONDARY RHEA2

Each socket supports up to 12 RDIMMs with ECC @6400MT/s.

The RDIMMs placed on the east and west side of each socket will share an I3C bus that will be exclusively connected to the socket (no connection to BMC).

Management and Debug

Figure 34 below is a description of the connections for board management within the [Rhea2-HPM], and with the DC-SCM (BMC board).

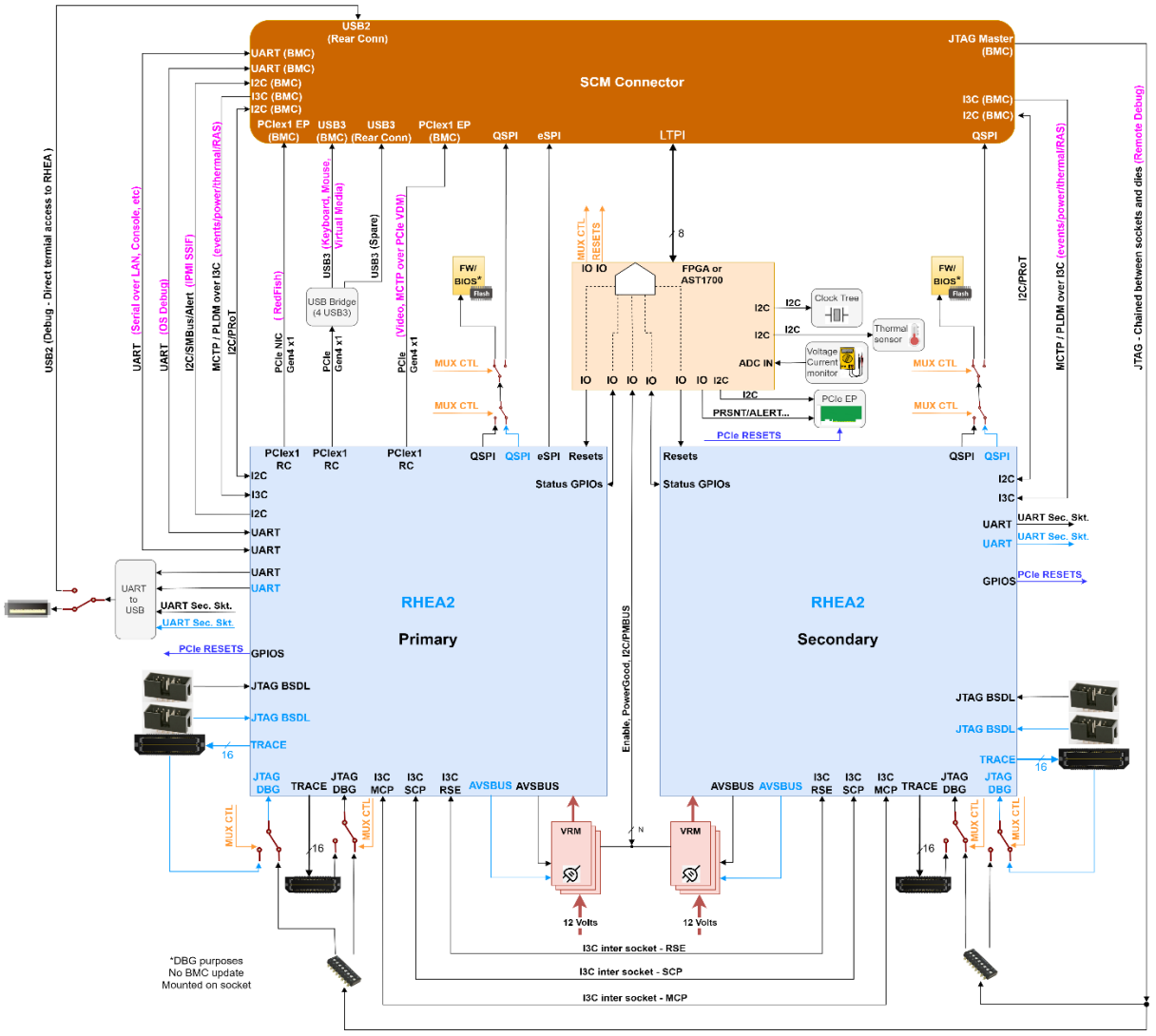

FIGURE 34 – DUAL SOCKET RHEA2-BASED HPM - BOARD MANAGEMENT INTERFACES

The following paragraphs will detail the schematic block of Figure 34.

DC-SCM connections

Board management of dual socket [Rhea2-HPM] complies with [SBMR] Level 4.

Following the Level M4 requirements of SBMR, below is an enumeration of the connections between the primary socket and the BMC on the DC-SCM board:

- 1x UART for serial over LAN, console, etc
- 1x UART for OS debug
- 1x I2C/SMBus/Alert for IPMI SSIF

- 1x MCTP / PLDM over I3C for events/power/thermal/RAS
- 1x I2C for Platform Root of Trust
- 1x PCIex1 Gen4 RC (SOC side) for a NIC interface (Redfish standard)
- 1x PCIex1 Gen4 RC (SOC side) for video over BMC
- 1x USB3.0 for keyboard, mouse, virtual media
- 1x JTAG for remote debug

Notes:

– JTAG of the secondary socket is chained with the JTAG of the primary socket and connected to BMC's master JTAG for remote debug.

Secondary socket has:

- An MCTP / PLDM over I3C interface to the BMC as well, to be able to boot secondary socket independently of primary for debug purposes
- An I2C interface for connection interactions with the platform root of trust located on the DC-SCM

The interfaces described above, and related to the SBMR specification, are represented on Figure 34 in **pink** colour.

In addition to the above interfaces, the following will be connected to the DC-SCM board:

- 2x QSPI interfaces (one per socket) to access the QSPI flashes on DC-SCM board for RHEA2 boot.
- 1x eSPI interface (for primary socket) to access to TPM on DC-SCM board.
- 1x USB2 interface to be connected to back panel for an UART to USB direct connection to both primary and secondary RHEA2 for debug purposes.
- 1x LTPI interface connected to the BMC chip to tunnel the following low speed signals (list is not exhaustive):
    - Primary socket resets and alert signals
    - Secondary socket resets and alert signals
    - PCIe endpoint low speed signals (alert, I2C, wake …)
    - On board thermal sensor I2C
    - On board clocks I2C
    - On board VRMs PMBUS
    - Enable and Power Good signals of on board VRMs
    - Various MUX select signals.

Onboard connections

Below is a description of the onboard signals that are not connected to the DC-SCM:

– 4x AVSBUS (one per die) for VRM telemetry and voltage control

– 3x I3C inter-die for inter-socket communication

– 2x JTAG debug connectors (one per socket, accessible from both dies)

– 4x parallel traces connectors (one per die)

– 4x JTAG BSDL connectors (one per die)

3.2.5 Software

The [Rhea2-HPM] will be delivered with a software stack including the firmware based on ARM Neoverse reference platform RDV3R1. Consequently, the deliverables names are based on ARM Neoverse processing elements:

- RSS Firmware: Boot-time & Run-time running on the RSS (Runtime Security Subsystem)
- SCP Firmware: Boot-time & Run-time running on the SCP (System Control Processor)
- MCP Firmware: Boot-time & Run-time running on the MCP (Manageability Control Processor)
- LCP Firmware: Boot-time & Run-time running on the LCP (Local Control Processor)
- AP Firmware: Boot-time & Run-time running on the Application Processors
- Any firmware related to chiplet subsystem (ex: PCIe, DDR…)

In addition to the above mentioned firmware, the [Rhea2-HPM] will be delivered with the following software stack running on Application processors:

- UEFI EDK2 BL33
- UEFI OS loader
- Linux Kernel
- Linux Distribution (busybox, buildroot, RHEL ARM Neoverse reference images)

The software will also incorporate the host-side drivers for ARM servers to enable interaction between the host-resident software stack and the accelerator's compute and memory resources, following the OpenMP task offload programming model. These drivers will be leveraged from the RISER project.

3.3 Single Socket Rhea2-based HPM

3.3.1 Requirements

According to [D2.1], the single-socket Rhea2-based HPM ([Rhea2-HPM]) is SDNO Class A335 compliant.

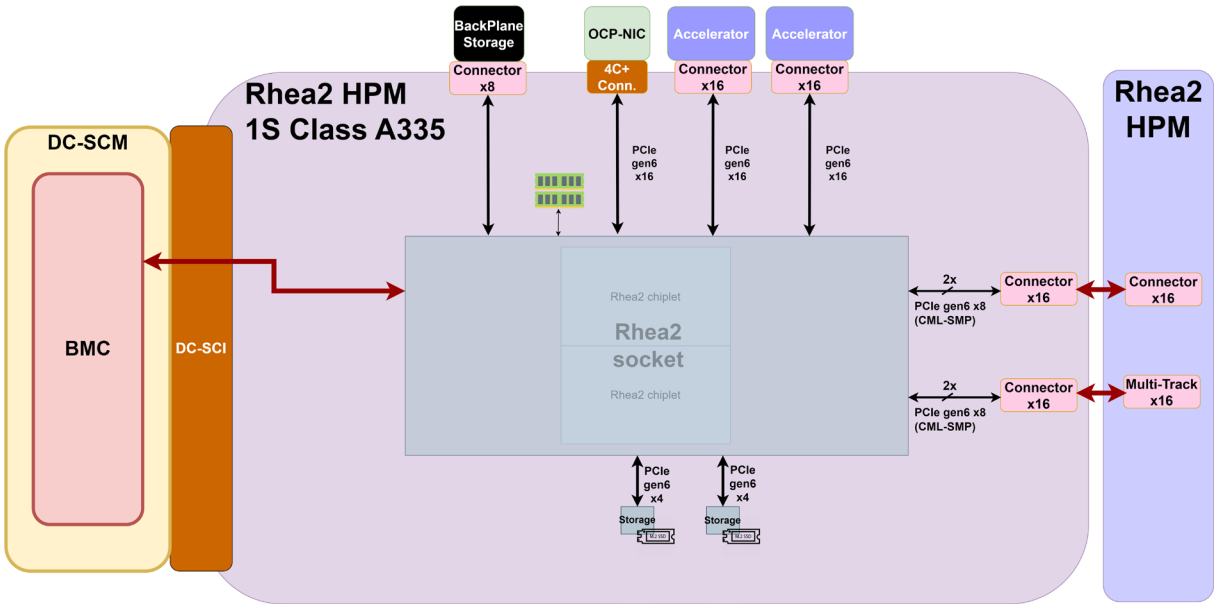

FIGURE 35 – SINGLE SOCKET RHEA2-BASED HPM - HIGH-LEVEL BLOCK DIAGRAM

3.3.2 Construction of the single socket [Rhea2-HPM]

The single socket board will have only the primary socket, which means only the primary socket's connections will be available.

4x connectors are added for the support of the 4x PCIe x8 Gen6 CML-SMP dedicated to all to all inter die communication (2 connected to primary socket and 2 connected to secondary socket).

Figure 36 shows the preliminary layout of the single socket Rhea2 based HPM.

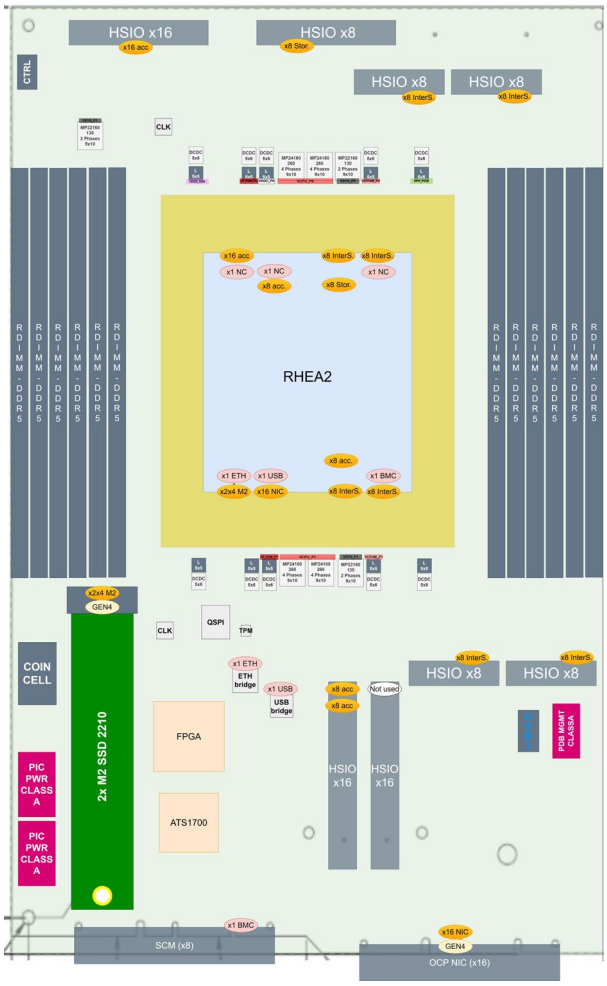

FIGURE 36 – SINGLE SOCKET RHEA2-BASED HPM - PRELIMINARY PLACEMENT

Listed below is a summary of the PCIe x16 functions:

- 2x PCIe x16 Gen6 connectors for accelerators
- 1x PCIe x8 Gen6 for storage back planes
- 4x PCIe x8 Gen6 CML-SMP for all to all inter-die communication
- 1x PCIe x16 Gen6 connectors for OCP-NIC
- 2x PCIe x4 Gen4 for M2 SSD connected to the primary socket
- 1x PCIe x1 Gen4 for USB3.0 bridge (UPD720201K8-701-BAC-A from Renesas)
- 1x PCIe x1 Gen4 for ETHERNET bridge (I225 from Intel)
- 1x PCIe x1 Gen4 for display port through DC-SCM's BMC
- 1x PCIe x1 Gen4 for Redfish protocol

Figure 39 below gives a more detailed view of how the connectors will be used for establishing the all to all connection between the 4 chiplets involved in a dual-HPM configuration:

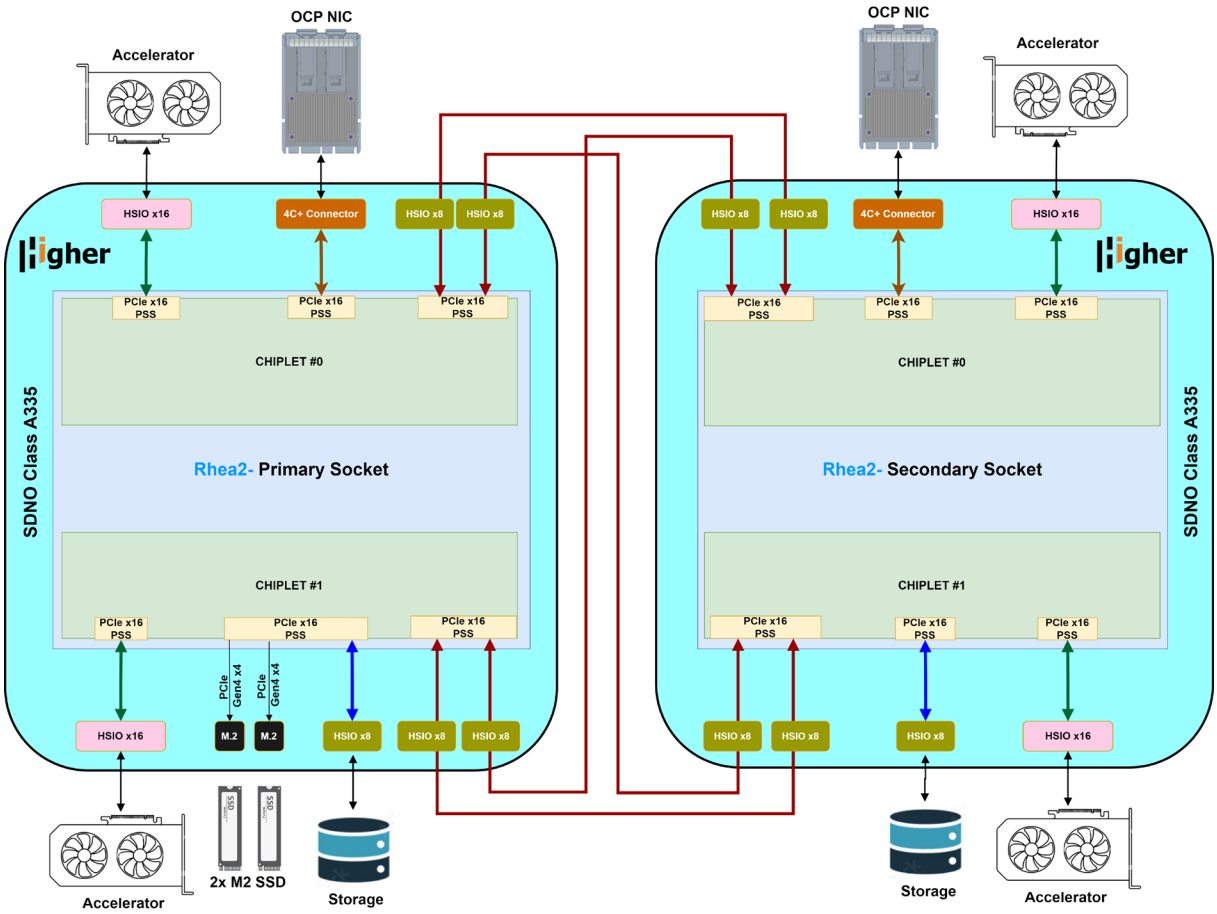

FIGURE 37 – SINGLE SOCKET RHEA2-BASED HPM - DUAL HPM CONFIGURATION

3.4 EPAC/EUPILOT Processor Module

The EUPILOT processor module will include the HIGHER carrier board, which can host up to two EUPILOT OCP Accelerator Modules (OAMs). The OAM complies with the OCP standard specification. Each OAM primarily supports one EUPILOT processor chip and four LPDDR4 memory devices. The carrier board provides high-speed external connectivity to both EUPILOT processor chips via the connectors of the two OAMs.

This connectivity includes high-speed links for direct OAM-to-OAM communication, allowing the EUPILOT chips to communicate directly, as well as PCIe links for interfacing with an external host or external PCIe devices when the EUPILOT chip is configured as a PCIe root complex.

To ensure signal integrity in high-speed communication, protocol-agnostic redrivers will be used for OAM-to-OAM links, and PCIe retimers will be employed for PCIe connections. Each OAM will utilize up to six MCIO SFF-TA-1016 connectors for inter-OAM communication and one MXIO SFF-TA-1033 connector for PCIe connectivity.

Figure 38 illustrates a high-level block diagram of the architecture, showing the carrier board providing high-speed connectivity to two EUPILOT OAMs.

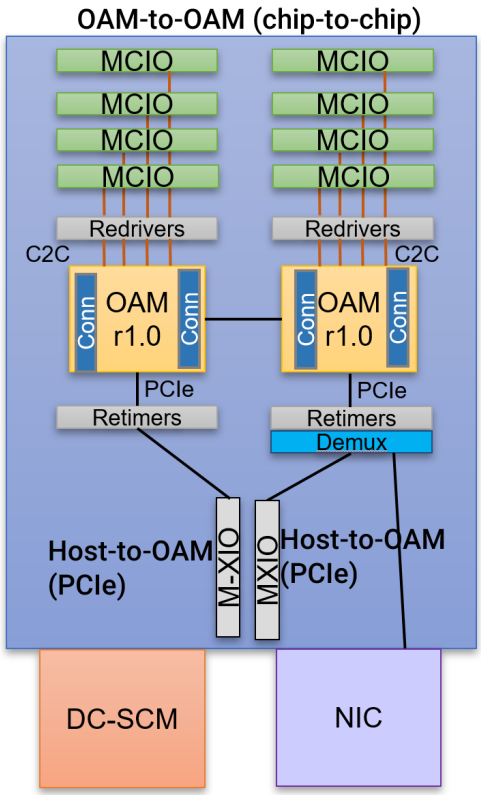

FIGURE 38 – HIGH-LEVEL BLOCK DIAGRAM OF EUPILOT PROCESSOR MODULE

Moreover, it is important to note that the carrier is deliberately designed to offer connectivity capabilities that exceed the baseline requirements of the EUPILOT OAM. The goal of this enhanced connectivity is to increase flexibility and enable broader exploitation opportunities. Specifically, the expanded connectivity is intended to support a wide range of OAMs, allowing the carrier board to be utilized in other designs and to accommodate future advancements and diverse application scenarios.

3.4.1 Hardware

3.4.1.1 EUPILOT OCP Accelerator Module

The EUPILOT OAM design will be provided by the EUPILOT project. The OAMs used within the HIGHER project will be manufactured and assembled as part of the HIGHER project activities. The OAM complies with the OCP OAM r1.0 v1.5 specification and operates with a 12V power input. Each OAM can consume up to 300W.

Figure 39 presents a block diagram of the OAM. At its centre is the EUPILOT chip, accompanied by four LPDDR4 memory chips. The module also integrates an AVR128 board management controller, which handles board power-up and status monitoring. The AVR is externally accessible via an I²C interface. Additionally, the board includes an SPI flash for booting the EUPILOT chip.

External high-speed connectivity consists of six chip-to-chip high-speed links that support a proprietary protocol, and one PCIe Gen5 link. Each chip-to-chip link comprises two lanes, while the PCIe link consists of four lanes.

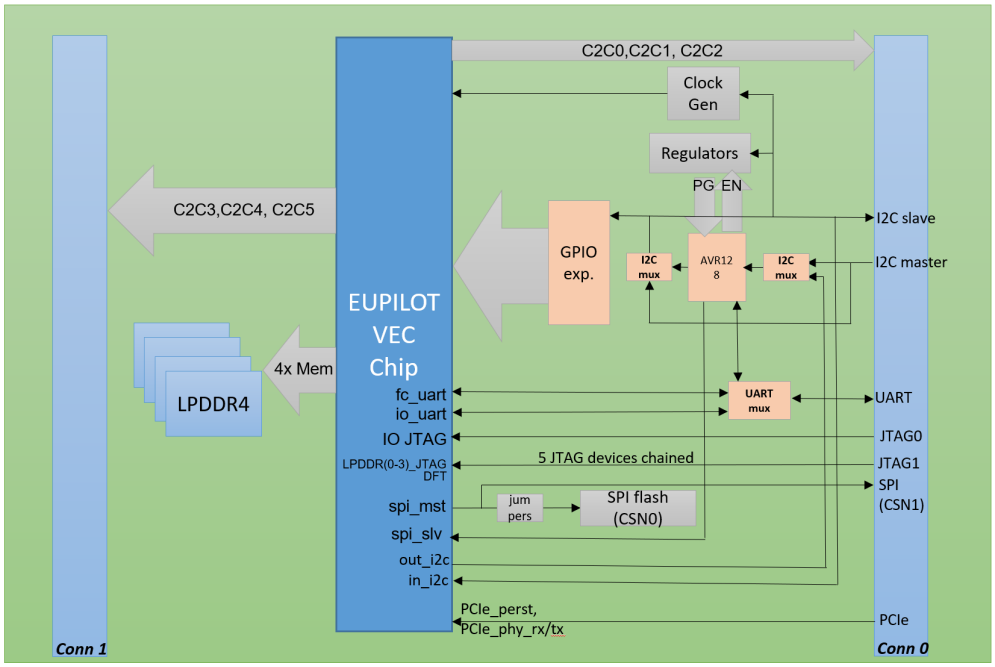

FIGURE 39 – BLOCK DIAGRAM OF EUPILOT OCP ACCELERATOR MODULE

1.1.1.1 HIGHER OAM carrier processor module

The carrier module complies with the OCP M-SDNO A335 standard as described in Figure 40. At the centre of the board, the carrier processor module hosts two OAMs. The specification defines two key zones near the top and bottom edges of the board, each with distinct features and functions.

In the M-SDNO A335 specification, the near-side zone is located closer to the reference datum of the board and allows for more flexible component placement. It typically includes a single or dual bi-directional PICPWR connector, which can be used for either power input or output, depending on system requirements. This zone supports varied chassis and cable routing configurations.

The far-side zone, positioned at the opposite edge of the board, has more rigidly defined connector placements. It may include one or two fixed-position power connectors designed for reliable blind-mate or cabled connections to the system infrastructure. The separation between these zones supports efficient power architecture planning and ensures compatibility across different board and chassis designs.

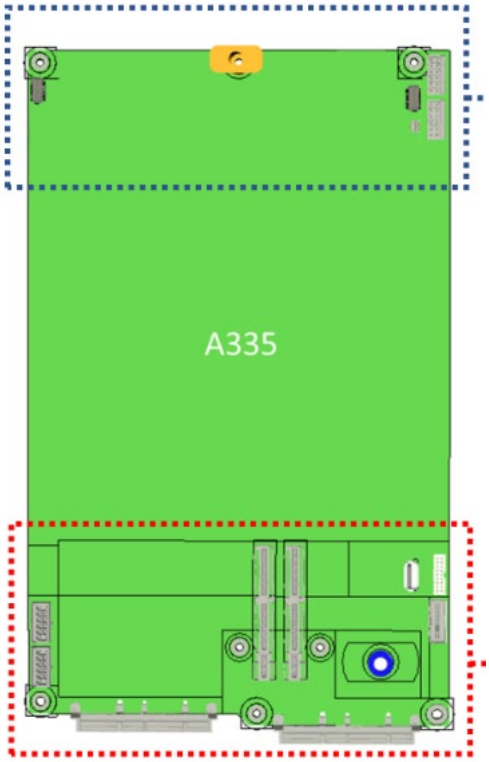

FIGURE 40 – OCP M-SDNO A335 FORM FACTOR

The Main characteristics of the carrier processor module are given in the table below:

OAM-to-OAM connectors	Up to six high-speed connectors
PCIe connectors	• 2x Connectors with PCIe x8 Gen6 RC PCIe links for storage
DC-SCM connector	• 2x UART • I2C • JTAG
NIC connectors	4x Connectors (2x6 + 12sb vertical header) 12V/12A per pin
Mechanical	M-SDNO A335

TABLE 2: EUPILOT CARRIER PROCESSOR MODULE MAIN CHARACTERISTICS

1.1.1.1 Component Placement

Figure 41 illustrates the initial placement of the primary components on the M-SDNO (Modular Scalable Disaggregated Network Offload) board. The layout accommodates two OCP OAM units, demonstrating compatibility with the M-SDNO form factor and highlighting the platform's support for high-density accelerator integration.

At the top side of the board, each OAM is connected via up to three high-speed connectors. These are designed to carry direct OAM-to-OAM data between the accelerators and other system components, such as host processors if required. Adjacent to these connectors is the 12V power input, which supplies the necessary low-voltage power for the OAM and peripheral circuitry located in the upper region of the board. The input voltage is typically stepped down on-board using power converters to drive 3.3V components requiring higher power density, such as clock generators and the OAM accelerators themselves.

At the centre of the board, four high-density connectors (two per OAM) are used to interface with the two OAMs. These connectors facilitate high-speed external connectivity to both EUPILOT processor chips and support the management of the OAMs. They adhere to the OCP OAM r1.0 v1.5 specification, ensuring standardized management and control interfaces. Each connector delivers a regulated 12 V power rail with a capacity of up to 300 W per module, thereby guaranteeing sufficient power delivery for high-performance accelerator operation.

On the bottom side, each OAM is interfaced with up to four additional high-speed connectors, which provide expanded bandwidth capacity for data-intensive workloads. Also located on this side is the 48V power input, which delivers higher voltage power.

Along the bottom edge of the board are connectors to the DC-SCM and a network interface card (NIC). The DC-SCM provides system control and management functions including platform security, while the NIC enables high-throughput networking for data ingress and egress.

To ensure signal integrity across these high-speed links - especially given the considerable trace lengths and potential for signal degradation at high frequencies - retimers or redrivers are integrated into the design. These components restore signal quality by compensating for losses due to impedance mis-matches, crosstalk, and other transmission effects inherent in high-speed board-level interconnects.

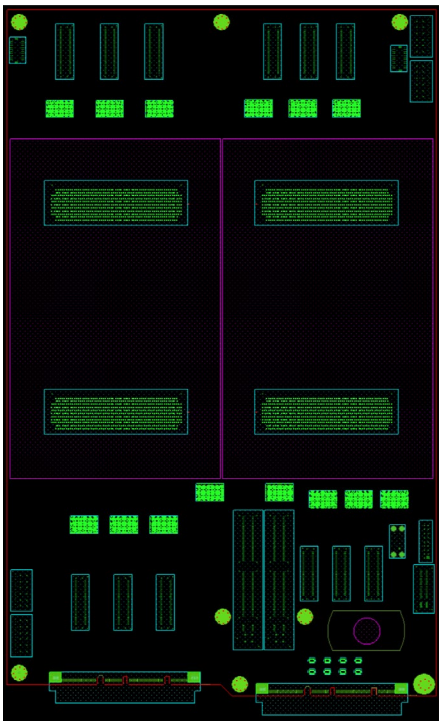

FIGURE 41 – INITIAL DRAFT PLACEMENT OF MAIN COMPONENTS ON EUPILOT PROCESSOR MODULE.

3.4.2 Management and Monitoring

The management and monitoring of the OAMs and the EUPILOT carrier board are handled by the BMCs located on the two OAMs and the BMC on the DC-SCM. Centralized control is provided by the BMC on the DC-SCM, which can communicate with the BMCs on both OAMs. Communication primarily uses I2C and UART interfaces. The BMC on the DC-SCM acts as the main management entity, responsible for overseeing all components on the carrier board and OAMs, as well as retrieving their status.

3.4.3 Firmware and Security

The EUPILOT carrier board is designed to interface with the DC-SCM platform through dedicated connectors, supporting a modular and standardized approach for system security. One of the primary objectives of the DC-SCM is to utilize its firmware in order to standardize platform security by physically decoupling and isolating critical components from the host system. The platform integrates advanced management and monitoring subsystems that facilitate secure control and oversight of the system's security state. At the core of the HIGHER DC-SCM framework is a dedicated hardware-based Root-of-Trust (RoT), which operates in isolation to ensure verified secure boot, firmware integrity verification, and platform attestation. The RoT enforces the validation of digitally signed firmware prior to the execution of any updates, thereby mitigating unauthorized code execution. Additionally, the platform integrates a TPM, which is employed to enable measured boot processes and support remote attestation by securing and reporting of cryptographic hashes of firmware components. This mechanism enables external entities to verify the integrity and authenticity of the system state prior to deployment or runtime operations.

3.5 Management Module

3.5.1 Make or Buy analysis

HIGHER aims to employ a DC-SCM that implements baseboard management controller (BMC) functionalities, i.e., remote server management, security, and control. It is worth noting that the most commonly deployed and uncontested BMC provider is ASPEED Technology Inc, a company from Taiwan.

HIGHER's original plan is to follow the OCP DC-SCM specification, to develop a BMC module in standardized form factor, applicable across diverse data center platforms.

In practical terms, this module was expected to embed a RISC-V processor with an FPGA (for flexibility in programmable I/O functionality), and offer a hardware-based Root-of-Trust (RoT) as the security foundation on which all sensitive procedures on the processor modules rely, including secure boot and remote attestation.

Along the work conducted over the past weeks, the HIGHER partners evaluated the complexity of implementing a full BMC + RoT on a single FPGA and the conclusion is that the base implementation of the BMC is high in expectation, especially the reached TRL with an FPGA implementation would not constitute the relevant ground for an alternative that would add value and enable competition with the existing implementation of Aspeed.

Nevertheless, the other conclusion is that there is likelihood of diversity for what concerns the root of trust, and the HIGHER partners rather focused on that part, landing on a solution where the plain DC-SCM will be based on Aspeed 2600 sourced from a supplier to be identified in subsequent tasks of the project, and the said DC-SCM will have the capability of hosting a daughter board dedicated to the Root of Trust that will be integrated on an FPGA.

3.5.2 RoT on a daughter card

Looking at solutions available on the market, HIGHER partners identified some use cases where the Root of Trust is not soldered directly on the PCB of the DC-SCM itself, but rather located on a daughter board.

Typically, the SCM202A from Flex (as referenced from the OCP market place) proposes such flexibility for what concerns the integration of the Root of Trust, as shown in Figure 42 and Figure 43 below:

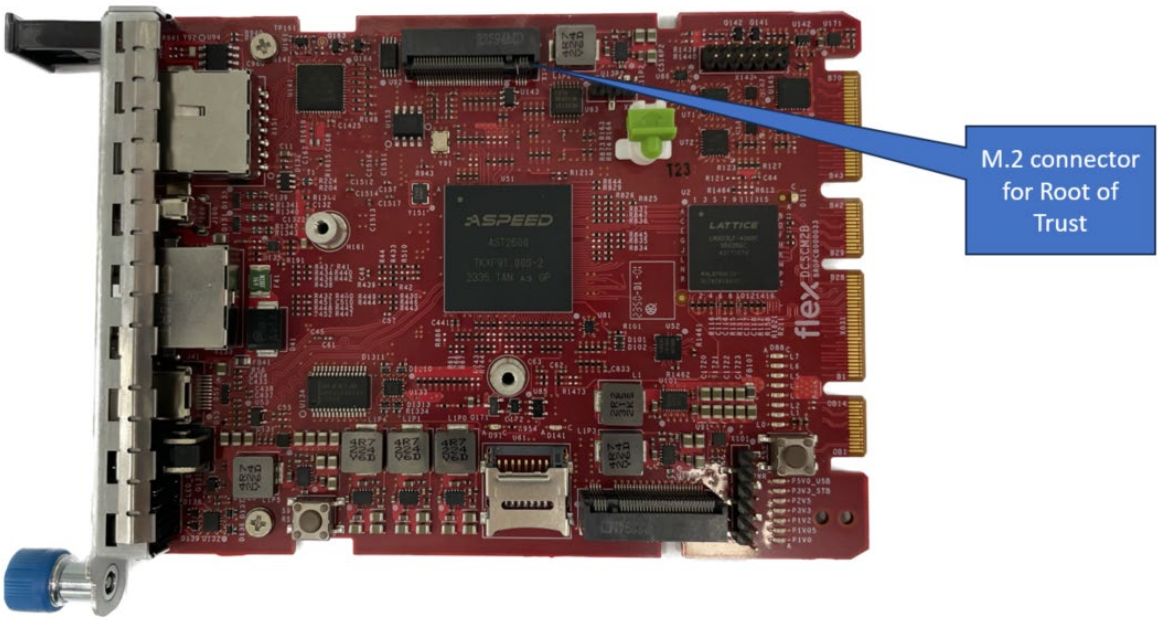

FIGURE 42 – FLEX DC-SCM - M.2 CONNECTOR FOR ROT

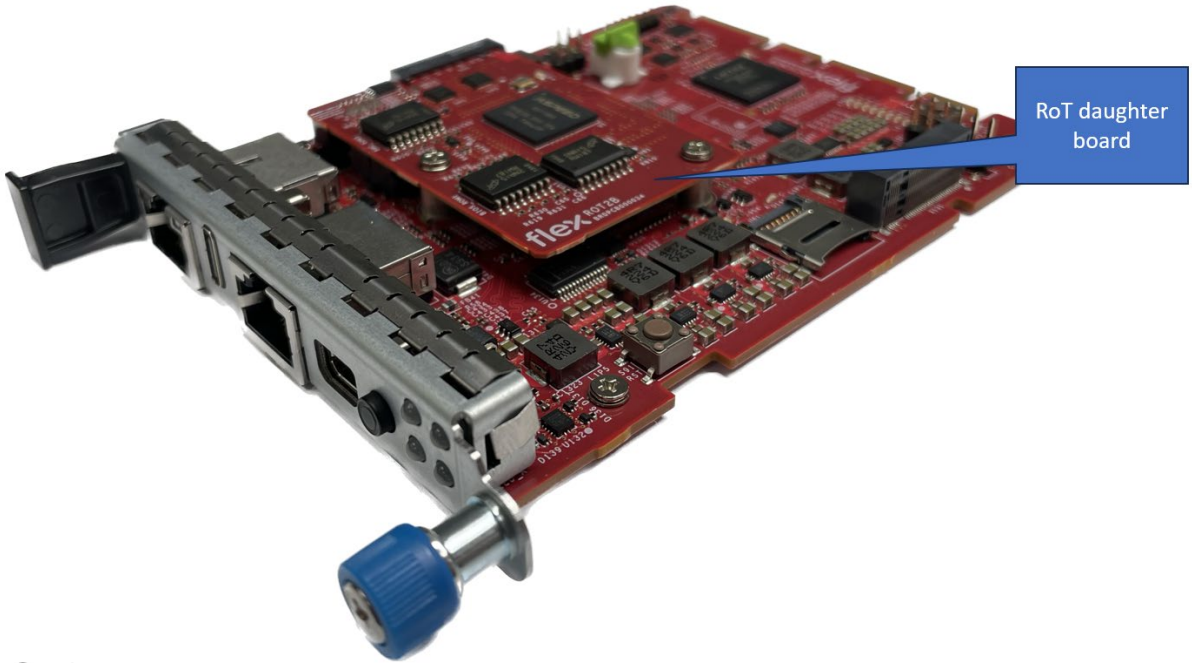

FIGURE 43 – FLEX DC-SCM - ROT DAUGHTER BOARD

AntMicro has a similar approach with their Artix DC-SCM (Artix DC-SCM).

The Figure 44 below shows a high level block diagram of the Artix DC-SCM with the root of trust:

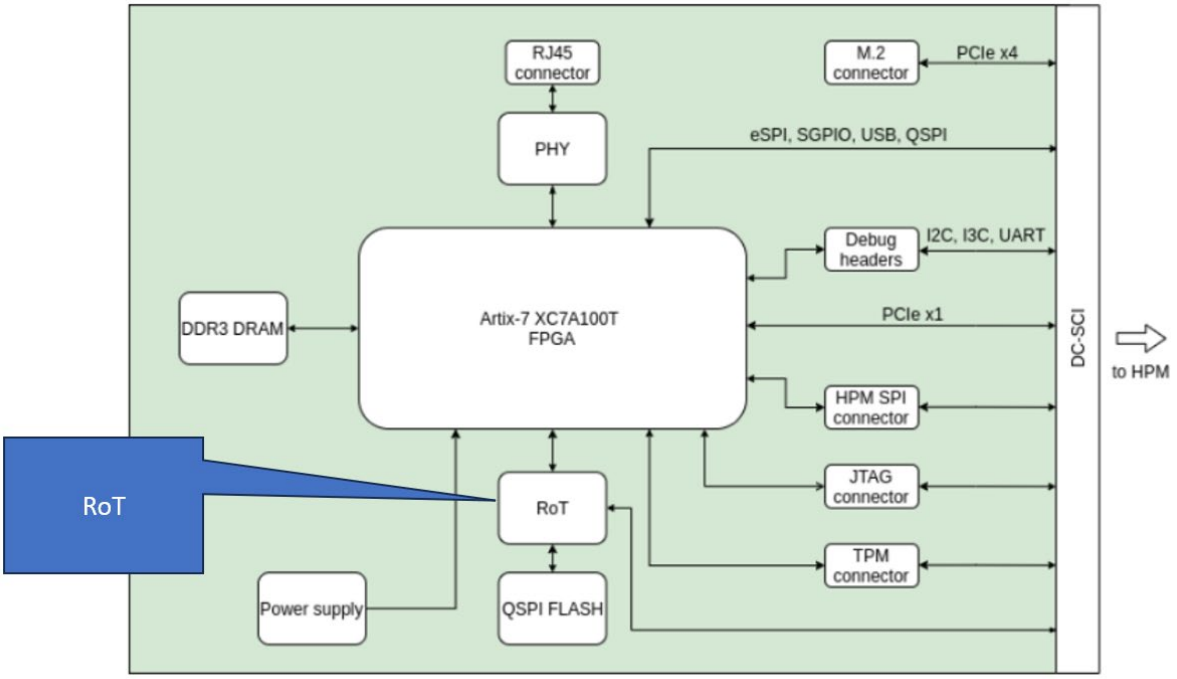

FIGURE 44 – ANTMICRO ARTIX DC-SCM - ROT ROLE

Figure 45 below details the mezzanine RoT module connector on the Artix DC-SCM.

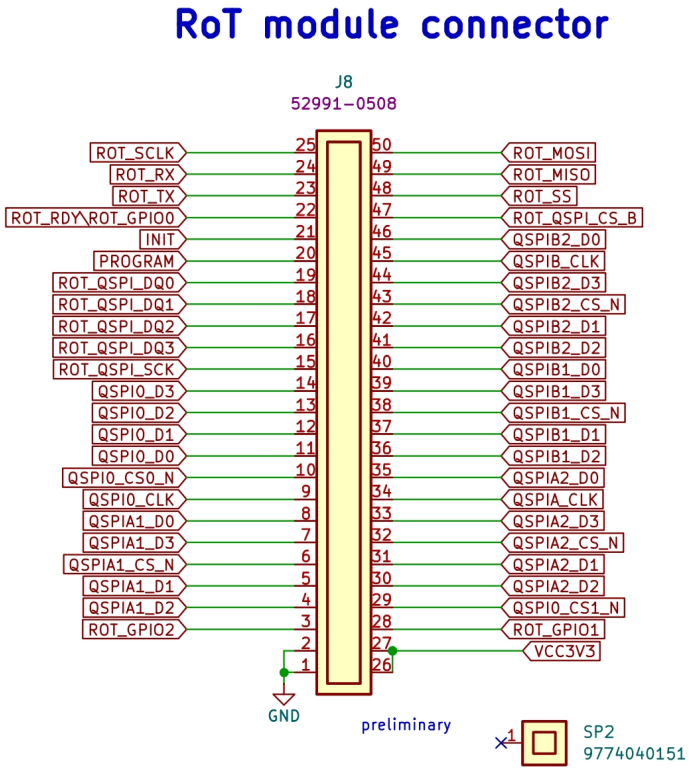

FIGURE 45 – ANTMICRO ARTIX DC-SCM - ROT MODULE CONNECTOR

These inputs as grabbed from the market are confirming the general impression that there is some opportunity for the introduction of a Root of Trust as daughter board, fitting a RoT interface to be defined from hardware and protocol interfaces, which could constitute a relevant contribution that the HIGHER project could potentially bring to the OCP project.

Figure 46 below is a high level description of where the RoT interface shall stand:

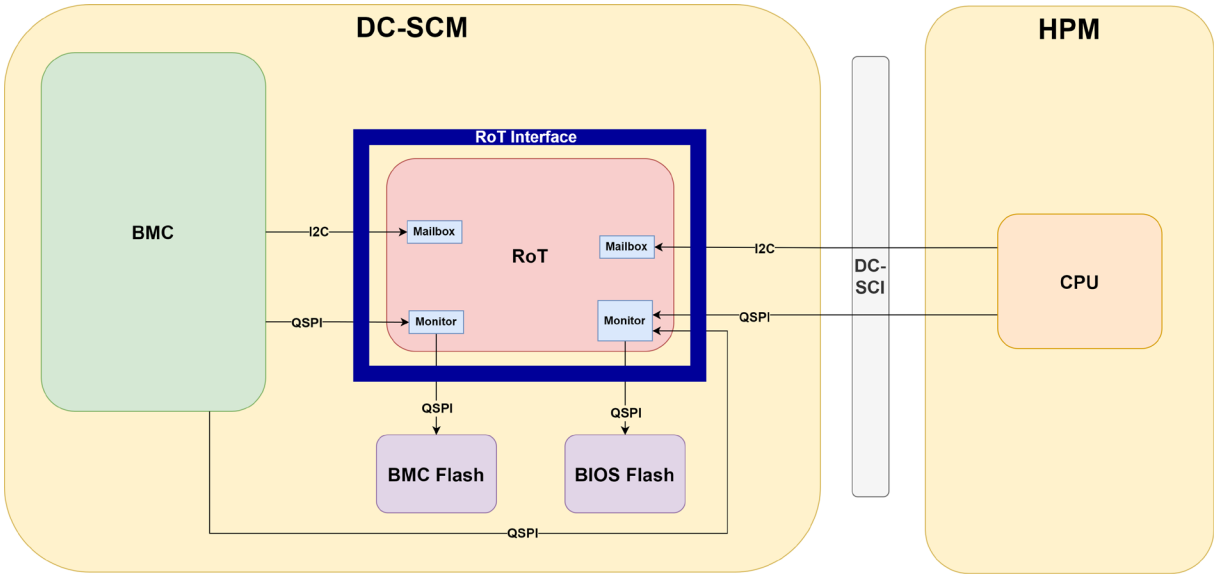

FIGURE 46 – ROT INTERFACE

3.5.3 Hardware

Figure 47 below is a block diagram describing the main interfaces of the required DC-SCM board:

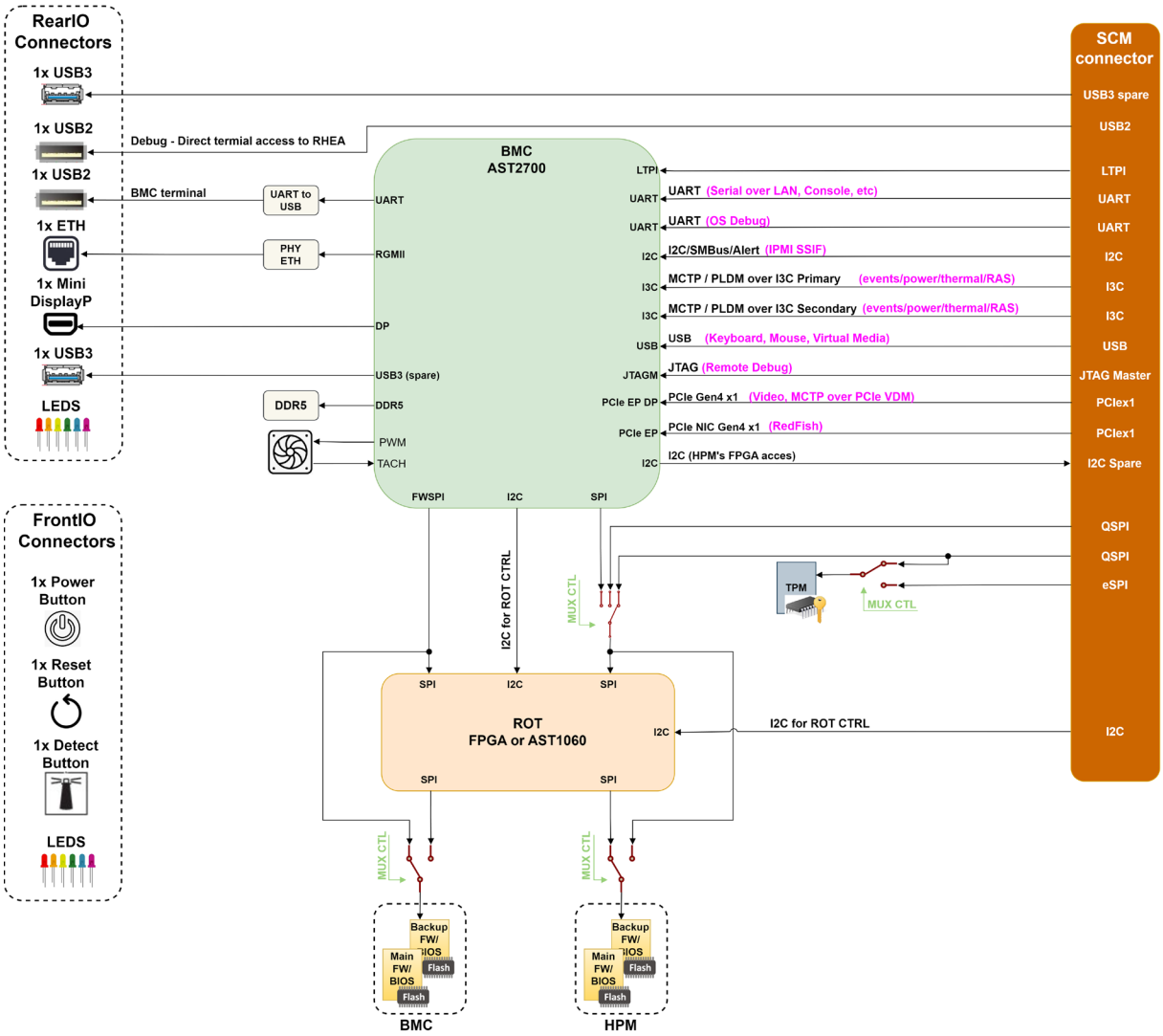

FIGURE 47 – DC-SCM - BLOCK DIAGRAM

Figure 47 lists the interfaces described in chapter 3.2.4, in the **Management Module** part.

In addition to that it describes the RoT and flash connections on DC-SCM board.

It also describes the front and back panel interfaces.

3.5.4 Description of BMC firmware

The BMC firmware will be based on OpenBMC.

Depending on the selected provider, the software adaptation will be performed by the provider or by HIGHER consortium members. The exact framework (e.g. JDM as proposed by Flex) will be decided during the WP3 activities.

3.5.5 RoT on FPGA

Recently, Lattice disclosed the innovative introduction of the Root of Trust for Measurement (RoTM) with Caliptra, and the seamless integration of these solutions into Field Programmable Gate Array (FPGA) technology implementations.

Lattice FPGAs incorporate “crypto agile” capabilities in Hardware RoT products to deliver future-proofed protection of server platforms and other connected device applications, such as PQC extensions. In addition, they provide a relevant platform for prototyping and testing new designs or features for Caliptra, which appears as a perfect match for HIGHER that will be exploited during the WP3 activities.

Reference: [Blog] Redefining the Root of Trust Architectures of Tomorrow.

3.5.6 Conclusion and next steps

As a conclusion:

- The plain DC-SCM will be sourced from a supplier to be identified. The potential suppliers identified are Pegatron, Flex and Mitac. In that list, the most advanced and constructive exchanges are established with Flex.
- The DC-SCM will be based on AST2600 (considering that AST2700 is not yet deployed in the market and none of the potential providers is proposing it).
- The DC-SCM shall have flexibility for integration of alternate Root of Trust such as Caliptra.
- HIGHER will work on the definition and specification of the interface between the DC-SCM and the RoT module.
- The RoT module is ideally a daughter card hosting an FPGA on which the RoT is integrated.
- HIGHER will explore the FPGA solutions from Caliptra.

3.6 Platform Root of Trust

The size and complexity of datacenters continue to increase and increasingly critical applications are trusted to those installations. In addition, the momentum of the Open Compute Project means that hardware components can be sourced from a variety of providers, which puts additional pressure on datacenter operators and users to ensure that the hardware and software components are trustworthy. Among the many threats that apply to such systems, supply chain attacks where some malicious components are sneaked into the machines have proved to be particularly damaging. In that context, a Root of Trust (RoT) is essential to provide a trust base, on which a chain of trust can be built to ascertain the trustworthiness of the application environment.

In modern servers, there are multiple Roots of Trust which are chained together. Their trustworthiness is assessed by a Platform Root of Trust (PRoT), which is typically placed on the server management module. Its first role is to ensure that the firmware used in the management module (e.g. Board Management Controller software and CPLD firmware) are genuine. Once the management module is up and running, the PRoT also provides services to the other components of the server (e.g. satellite controller, main CPU, network interfaces), typically to ascertain that the RoT placed in those components are also trustworthy and vice versa, and to prove to those components that the management module is to be trusted.

There are quite a few RoT solutions from a diversity of vendors in the marketplace. However, the Caliptra approach[1] has drawn a lot of attention lately. Indeed, its specification is public (whereas most alternatives are vendor-specific) and has been incepted by the Open Compute Project and is maintained by the Chips Alliance. In addition, there is an open-source reference implementation made available by the Chips Alliance, a series of projects under the Linux Foundation. Furthermore, an important differentiator between Caliptra and other RoT solutions is that Caliptra is intended to be used as a Silicon Root of Trust, directly integrated into BMCs or SoCs, whereas traditionally the RoT was implemented in a separate chip communicating with the processors using low-speed protocols such as I2C. Integrating the RoT into the processor chip renders physical attacks much more challenging. It is also worth noting that there are already 2 versions of the Caliptra specification. While Caliptra specification 1.0 describes a feature-complete Silicon Root of Trust, Caliptra specification 2.0 adds support for Post-Quantum Cryptography.

Within HIGHER, we aim to use Caliptra to secure both the Rhea2 and EPAC/EUPILOT HPMs. As discussed in Section 3.5, the Management Module will provide a Caliptra Platform Root of Trust. The most likely implementation will leverage Lattice Mach CPLDs devices. In addition, an emulation environment based on QEMU will be used to get developments started ahead of hardware availability. In both cases, we will focus on implementing workflows that support requirements described in deliverable [D2.1]. Hence, we may decide to work with Caliptra Specification 1.0 if this accelerates the implementation. The following sub-sections describe the workflows that will be implemented in the project.

3.6.1 Management module integrity

Using the relevant Caliptra APIs, the management module will be measured upon reset and the initial firmware binaries will be authenticated before their execution on the management module chips. This will take advantage of Caliptra Root of Trust for Measurability (RoTM) capabilities. The goal is that any mutable component used in the management module is verified before its use. This includes the bitstream of all CPLD and FPGAs and software executed on the different processing cores of the board. This workflow will fulfil requirement UCR-SYS-RT-76400.

[1] https://github.com/chipsalliance/Caliptra

3.6.2 Secure management module update

We will generate multiple updates to the mutable components of the management module, and ensure that properly authenticated binaries will be allowed, whilst altered binaries will not. This workflow will fulfil requirement UCR-SYS-MM-76201.

3.6.3 A/B firmware updates

We will ensure that the selected management module supports A/B firmware updates, where two versions of the firmware are stored at any time on the management module, so that in the event of an update failure (e.g. Binary is corrupted or update process is interrupted), former firmware can still be used. This workflow will fulfil requirement UCR-SYS-RT-76404.

3.6.4 HPM integrity

When the reset of the Host Processor Module is released, the management controller placed on that board, being satellite controller or microcontrollers in the main CPU, will initiate contact with the Platform Root of Trust to ensure that the management module is trustworthy and provide guarantees to the management module that the HPM integrity is preserved. This workflow will fulfil requirement UCR-SYS-RT-76401.

3.6.5 Transfer of ownership

We will perform a scenario where an entity A transfers the ownership of the server to an entity B, and ensure that (i) before the transfer only entity A can successfully provide updates to the management module, and (ii) following this transfer, only entity B is able to provide updates. In practice, this is necessary to make sure that solely the current owner is able to alter the management module, from the board manufacturing until the final deployment. This workflow will fulfil requirement UCR-SYS-RT-76402.

3.6.6 Defence against rollback

We will verify that former authenticated firmware will not be allowed to execute in the management module. In effect, in case a security flaw has been discovered and fixed in an updated version, an attacker may try to rollback to the former unsecure version. This workflow will fulfil requirement UCR-SYS-RT-76403.

3.7 CXL memory disaggregation

The Compute eXpress Link (CXL) is an open standard created and updated by industrial entities across the globe. CXL defines an interconnection protocol for devices, such as CPUs, GPUs, NICs, and FPGAs, and enables coherency and memory semantics among them. CXL leverages PCIe hardware links, therefore making a well-suited technology for disaggregated data centres. CXL defines three sub-protocols, namely cxl.io (device configuration), cxl.cache (coherency between devices) and cxl.mem (atomic operations for accessing memory). Moreover it defines three device types, namely type 1 (accelerators with optional caches using cxl.io and cxl.cache), type2 (accelerators with caches and expanded memory using cxl.io, cxl.mem and cxl.cache), and type 3 (memory expanders using cxl.io and cxl.mem). D2.1 has already listed the CXL memory disaggregation requirements; it should be noted that CXL-based memory expansion will be enabled only by the Rhea2 HPM, since it exposes CXL-based interfaces.

3.7.1 CXL memory Pool Manager (CPM) architecture

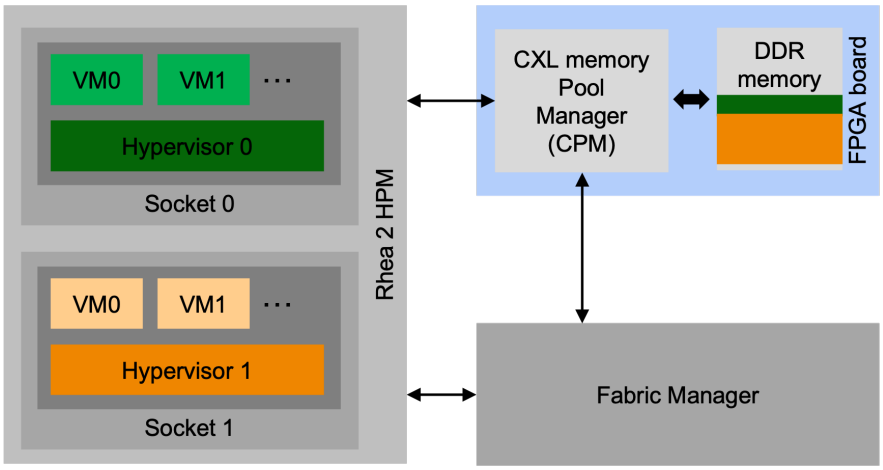

FIGURE 48 – LOGICAL INTERCONNECTION OF THE FABRIC MANAGER, RHEA2 HPM AND CXL MEMORY POOL

Figure 48 shows the logical interconnection between the Fabric Manager (FM), Rhea2 HPM and the CXL memory Pool Manager (CPM). The FM, which is executed in the BMC, is responsible for configuring the memory pool ownership for the hypervisors running on each socket, based on the workload memory requirements. At boot time, the CMP will be visible as a type-3 [CXL] Multi-Logical Device (MLD), with each memory slice assigned a Logical Device (LD) ID. The total memory space will be visible (but not enabled) to both hypervisors, and the FM will statically enable memory slices for each hypervisor.

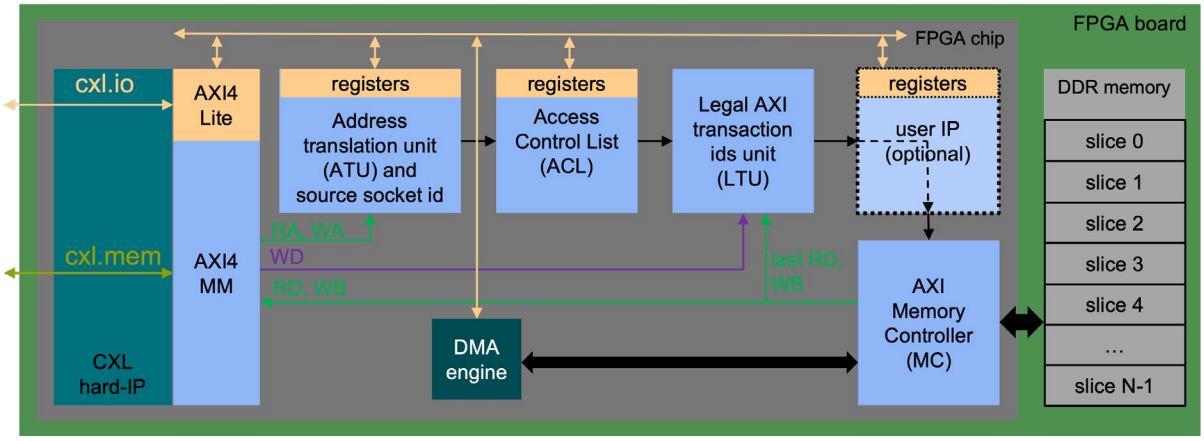

FIGURE 49 – CXL MEMORY POOL MANAGER (CPM) ARCHITECTURE

Figure 49 shows the CPM internal architecture that will be implemented to an FPGA card where the FPGA chip is connected to DDR4 memory. The FPGA will integrate a hard-IP that implements the CXL protocol, compatible with versions 1.1, 2.0 and 3.0. The CPM will be exposed as a type-3 CXL device; it can be configured either as Multi-Logical Device (MLD), where each memory slice is assigned a Logic Device ID (LD-ID), or a Single-Logical Device (SLD). It will use the cxl.io protocol for configuration commands and status updates, and the cxl.mem protocol for load / store operations from the Rhea2 HPM to access the DDR memory. As such, the CXL hard-IP exposes AXI4-Memory Mapped (MM) and AXI4-lite compatible interfaces to the programmable logic to forward cxl.mem and cxl.io requests respectively to the user logic.

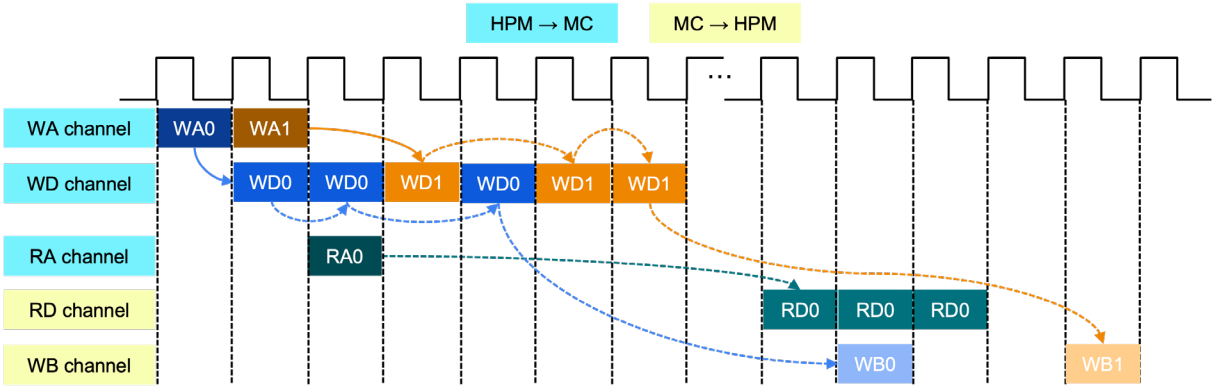

FIGURE 50 – AXI4 MM PROTOCOL TRANSACTIONS FLOW

Before diving into the details of the CPM architecture, we provide a quick description of the AXI4-MM protocol transactions. As shown in Figure 50, the AXI4-MM protocol supports 5 type of transactions, namely Write Address (WA), Write Data (WD), Read Address (RA), Read Data (RD), and Write Response (WB). Each WA transaction is accompanied by WD ones that carry the data to be sent from the HPM to the memory controller (MC), and when this is done, the MC sends back to the HPM a WB transaction to acknowledge the write outcome (success or error). Finally, RA transactions are sent from the HPM to the MC, which returns RD transactions that carry the data read from the memory.

Following the architecture depicted in Figure 49, the CPM comprises the following pipeline of operations towards checking which memory requests are legal before accessing the DDR memory:

- An Address Translation Unit (ATU) that converts HPM Physical Addresses (HPAs) to Memory Pool Addresses (MPAs), and extracts also the source socket id that generated a memory request. It includes a register file for address translation configuration (e.g. offset between HPA and

MPA, and slice granularity), which can be accessed by the AXI4-lite interface, as instructed by the FM.

- An Access Control List (ACL) that stores information related to memory slice ownership. It includes a register file for adding / removing slices to a socket, as well as updating access permissions, as instructed by the FM.
- A Legal AXI Transaction ids Unit (LTU) that caches ids of legal transient AXI transactions, i.e. transactions that are validated for accessing the requested memory area.
- An AXI memory controller that interfaces the external DDR memory.
- A DMA engine that enables fast slice duplication, if requested by the FM.
- Optionally, a user IP for in-situ data processing before accessing the memory. The IP includes a register file for configuration and status monitor.

The CPM logic forwards ingress WA and RA requests to the ATU to extract the address range to be accessed, apply the HPA-to-MPA translation, and finally read the socket id that issued the transaction. The HPA and socket id are forwarded to the ACL, which checks if the transaction is valid in terms of slice ownership and access permission, and forwards the result along with the transaction id to the LTU.

The LTU caches ids of transient RA and WA transactions, which are associated with their corresponding RD and WD ones. As mentioned, when RA transactions are processed, read data from memory are returned in the form of consecutive RD transactions, whereas WA transactions are followed by multiple WD ones that "carry" the data to be stored. Therefore, legal RAs / WAs plus their WDs are forwarded to the memory controller. As soon as RAs are processed by the MC, the latter will generate its RDs that contain the read data once they are returned from the memory, whereas for WAs+WDs the memory controller will generate its WB transaction (once they are written to the memory) to signal the HPM that the store command is done. Therefore, WBs and the last RD that is associated with an RA, are also sent to the LTU, which removes their ids from the list of legal transient ids. However, in case of illegal RAs and WAs, the LTU masks the MPA address with an illegal one to create a failure during address decoding, thus forcing a "DECERR" response for unauthorised WA, RA transactions.

3.7.2 UC4 example

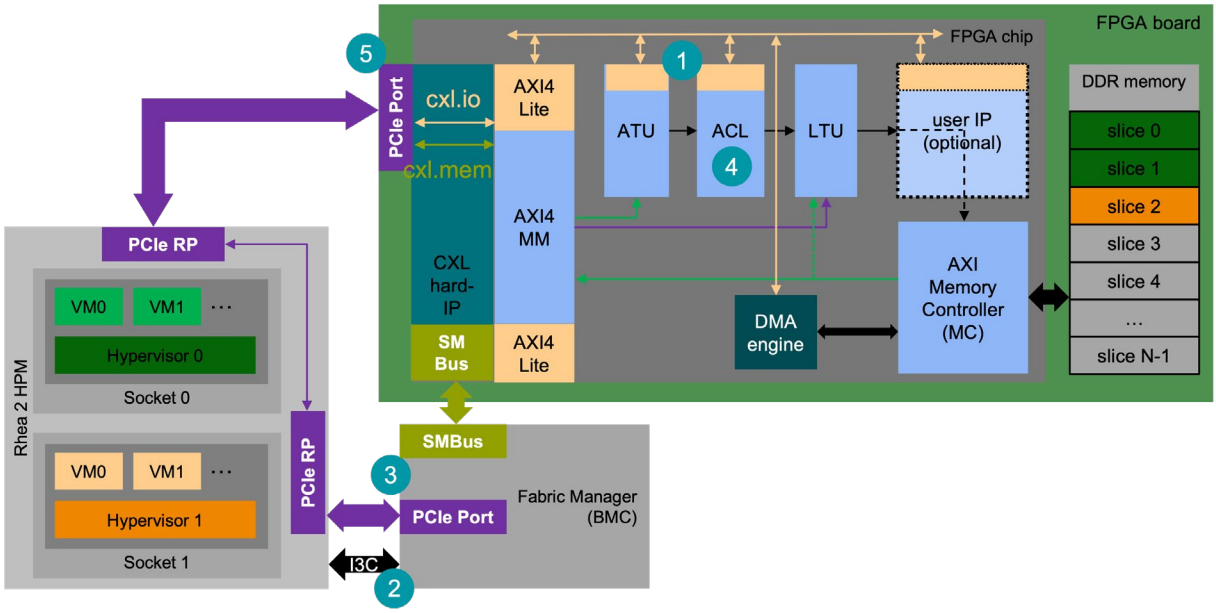

FIGURE 51 – CPM CONFIGURATION AND USAGE BY THE HPM

Figure 51 shows the options for connecting the CPM with the FM; the latter may use either an SMBus-based direct connection with the CPM or use the Rhea2 PCIe tree. As an example, these are the required steps so the FM can configure the CPM and then allocate additional memory for the Rhea2 HPM:

1. The FM assigns LD-IDs to each memory slice (in case of MLD configuration) or a single LD-ID in case the CPM is configured as a Single-Logical Device, and clears the ACL and LTU modules. Moreover, the FM configures the address offset that will be applied to each physical address from socket 0 and socket 1.
2. Hypervisor 0 (H0) and Hypervisor 1 (H1) request from the FM to allocate 2 and 1 slices respectively from the memory pool.
3. The FM sends to the CPM the hypervisor id and number of slices to be allocated for each socket;
4. The CPM updates the ACL with 2 entries; the 1st entry designates that H0 has access to slices 0 and 1, whereas H1 has access to slice 2.
5. The CPM informs the FM that slice allocation is done, and the FM informs H0 and H1 that a new memory space is now available for hot-plugging.

At this point, both H0 and H1 can use the additional memory slices for their VMs. Based on future workload requirements, H0 and / or H1 may ask the FM to free the allocated slices from the memory pool. In this case, the FM will again send an unbind command to the CPM, which will remove the required metadata from the ACL. When this is done, the CPM will notify the FM, which will instruct the hypervisor to unplug the previously hot-plugged memory space.

3.8 Associated System Software

In this document, the specification for the HIGHER systems is provided, and is mostly focused on hardware components. Firmware and software components are also presented, as they are required to enable the desired functionalities. In most cases, those components are presented along with the corresponding hardware. But there are also software components that are critical to meet the requirements defined in [D2.1]. Hence, the following section provides a brief list of both types of components, to provide a consistent overview of the System Software required to reach the envisioned HIGHER systems functionalities. Further details will be provided by deliverables produced by the WP4.

3.8.1 Emulation environments

In order to start the developments of the System Software in WP4, several emulation environments will be used. They will rely on QEMU as well as domain-specific emulators. When needed, those emulators will be modified for the purpose of the HIGHER project. For instance, an emulation platform has been set up to emulate Caliptra, and enable the related developments before the availability of the hardware.

3.8.2 Management module core firmware

The software running on the Board Management Controller (BMC) - which controls the Management module - will be tailored to the Host Processor Modules (HPMs) designed in HIGHER. To that end, OpenBMC will be customized to match the exact Data Center Secure Control Interface (DC-SCI) used by those HPMs, and to provide common services such as power cycling, server monitoring and secure updates (c.f. Section 3.6).

3.8.3 Management module services

In addition to commonly present services, the management module will support use-case specific features. First, in cases where an HPM is connected to another HPM through high speed cables (e.g. to offload computations from one to the other), the BMC will provide the necessary means to configure both sides adequately. Second, to support the CXL memory disaggregation use case described in Section 3.7, the necessary Fabric Management (FM) features will be implemented in the BMC.

3.8.4 Rhea2 HPM boot firmware

The various firmware elements required to complete the boot of the Rhea2 HPM will be produced. This includes binary files used by the HPM satellite controller and the main ARM CPU boot flow (e.g. BIOS). In addition, those components will support a secure boot flow, in cooperation with the management module. Furthermore, the discovery of the OCP-NIC and CXL devices will be implemented.

3.8.5 EPAC/EUPILOT boot firmware

The various firmware elements required to complete the boot of the EPAC/EUPILOT HPM will be produced. This includes binary files by the RISC-V CPU boot flow (e.g., BootROM and OpenSBI). In addition, those components will support a secure boot flow, in cooperation with the management module.

3.8.6 Linux OS images

HIGHER will provide Linux images for all HIGHER platforms. Further details will be provided in WP4.

3.8.7 Device drivers

Device drivers necessary for both x86 and Rhea2 HPMs to offload workloads onto the EPAC/EUPILOT processors will be enabled/developed. Those drivers will allow one to send a binary executable and data over to those processors, execute this binary, and retrieve resulting data. This is required by UCR-SW-DIST-54101.

3.8.8 OpenMP runtime

On top of device drivers described above, the OpenMP support for the EPAC/EUPILOT processors will be enabled/developed. This is required by UCR-SW-DIST-54100.

3.8.9 MPI runtime

In order to support HPC workload, HIGHER will provide an MPI implementation for both the Rhea2 and the EPAC/EUPILOT HPM platforms. This is required by UCR-SW-DIST-54202.

3.8.10 ML/AI libraries

HIGHER will provide libraries to support the ML/AI use cases for both the Rhea2 and the EPAC/EUPILOT platforms. Those are required by UCR-UC-PAAS-54300 and UCR-UC-PAAS-54300 respectively.

3.8.11 MetaOS, ColonyOS and resource discovery

ColonyOS is a distributed meta operating system (MetaOS) that enables orchestration of computational workloads for a diverse set of applications. It supports a continuum of execution environments that spans edge, cloud, and high-performance computing systems. ColonyOS consists of colonies servers that provide OS-like functionalities such as processes and a file system, and a network of executors that carry out computational jobs that are submitted to the colonies servers.

The Colonies servers are dependent on having a host operating system such as Linux, macOS, or Windows. Additionally, they use TimescaleDB to store information about system resources, executors, and a ledger of the execution history. By contrast, executors can have different dependencies based on the type of executor. The required base functionality for an executor to communicate with a Colonies server is limited to a JSON-based REST API using end-to-end encryption, but the executors may be integrated with larger frameworks such as Kubernetes or Slurm.

We expect to use Linux as our main OS for ColonyOS development within the HIGHER project.

4 Make or Buy Analysis

The make or buy analysis has focused on the DC-SCM and is extensively described in chapter 3.5.1.

For what concerns the OCP-NIC, relevant units will be procured in the subsequent phases of the project. There is no relevance to consider HIGHER making a custom OCP-NIC as these components are generally available and are not considered as critical in terms of sovereignty.

2CRSi has already provided a reference for an OCP-NIC model: BCM957416N4160C from Broadcom.

FIGURE 52 – OCP-NIC EXAMPLE

However, a thorough inventory of the candidate models will be conducted in the subsequent phases of the project, when the constituent of the servers, as described in chapter 3.1, will be sourced.

5 Typical cloud infrastructures

The two innovative compute platforms being developed as part of the HIGHER project - a single/dual-socket ARM-based Rhea2 Host Processor Module (HPM) and a RISC-V-based EPAC/EUPILOT accelerator module - can be integrated effectively into a modern cloud provider's infrastructure. Designed in line with current physical and software standards, these platforms are well suited for deployment within CloudSigma and other cloud providers environments, supporting a broad range of use cases from general-purpose compute to specialized acceleration.

5.1 Physical Integration

As the ARM and the RISC-V platforms are built according to Open Compute Project (OCP) standards, this ensures compatibility with contemporary data center deployments. The ARM platform, offered in SDNO Class A and C, is physically compatible with standard 19-inch rack systems, supporting both high-density and traditional server layouts. The EPAC/EUPILOT module integrates via the OCP M-SDNO A335 carrier board, supporting up to two OCP Accelerator Modules (OAMs). It also conforms to standard rack and interconnect specifications, including MCIO (SFF-TA-1016) and MXIO (SFF-TA-1033) connectors. Figure 53 below illustrates the physical integration.

Power and thermal design adhere to OCP norms, with the Rhea2 module utilizing four 12V/12A connectors and the EPAC/EUPILOT modules supporting up to 300W per OAM. This alignment allows these systems to be deployed directly alongside existing x86-based servers without requiring changes to rack infrastructure, cabling, or facility cooling design.

FIGURE 53 – PHYSICAL INTEGRATION OF ARM RHEA2 HPM AND RISC-V EPAC/EUPILOT ACCELERATOR MODULE IN TYPICAL CLOUD INFRASTRUCTURES.

5.2 Software Integration

Both platforms are designed for cloud-native and high-performance workloads and can be integrated into our existing software stack:

Compute and Virtualization: The ARM-based Rhea2 HPM supports full virtualization (e.g., KVM/QEMU), making it suitable for multi-tenant environments and traditional compute roles. With DDR5 memory support (up to 128 GB per RDIMM) and PCIe Gen6 connectivity, it is well suited for general-purpose compute, high-performance computing (HPC), and AI/ML workloads. The RISC-V-based EPAC/EUPILOT module is optimized as an accelerator and supports execution of compute kernels via OpenMP and MPI, complementing traditional compute nodes in heterogeneous deployments.

Containerization and Orchestration: Both modules support mainstream Linux distributions and container runtimes, allowing for integration into Kubernetes-based orchestration platforms. ARM already is fully compatible with Kubernetes and for RISC-V, k3s has already been ported and is under steady development. The Rhea2 module's robust PCIe layout (e.g., 4x PCIe x16 Gen6 for accelerators, 2x PCIe x8 Gen6 for storage) and the high-speed interconnects of the EPAC/EUPILOT accelerator support efficient containerized deployments for microservices and batch workloads alike.

Object Storage and CXL Memory Disaggregation: Both modules can serve as object storage nodes, handling large-scale unstructured data for cloud storage solutions. Additionally, the Rhea2 HPM supports CXL memory disaggregation, allowing dynamic memory allocation across sockets via the CXL Memory Pool Manager (CPM). This capability enhances resource utilization for memory-intensive workloads, such as big data analytics, aligning with the M-FLW use case.

Security and Management: Integration with OCP-compliant DC-SCM modules provides secure boot via the Caliptra Platform Root of Trust (PRoT), and remote management through OpenBMC and Fabric Manager (FM), aligning with data center operational requirements.

Figure54 below illustrates these main software layers for integration.

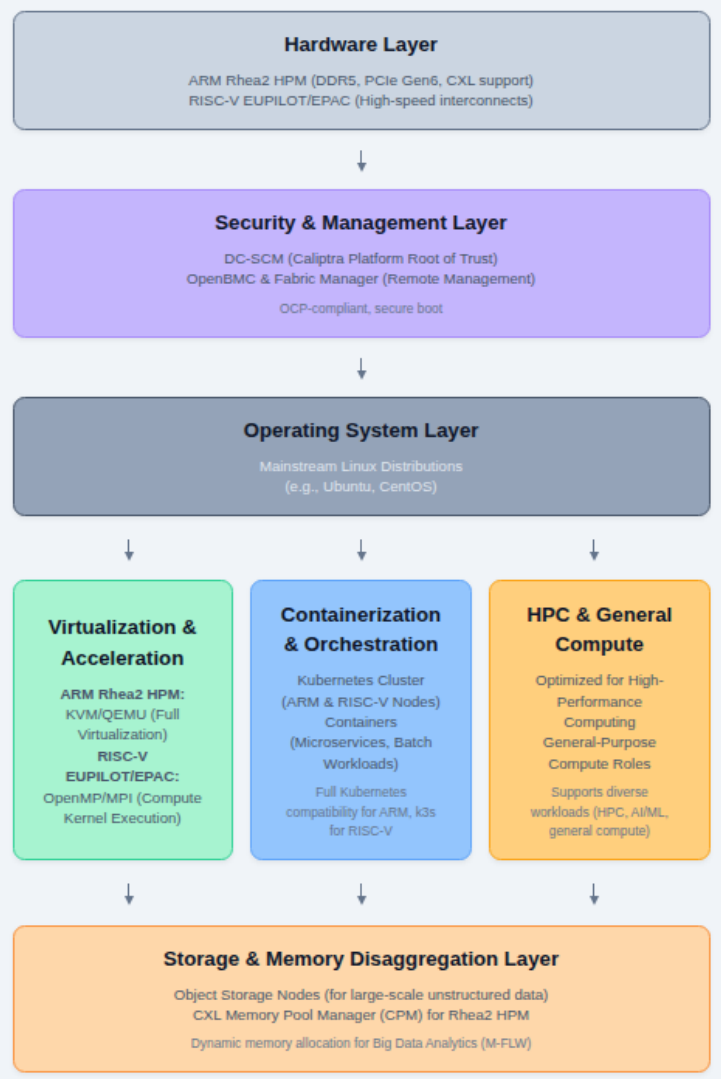

FIGURE 54 – MAIN LAYERS IN HIGHER SOFTWARE STACK TO SUPPORT CLOUD-NATIVE AND HIGH-PERFORMANCE WORKLOADS

5.3 Alignment with CloudSigma Infrastructure Use Cases

The Rhea2 HPM's performance, virtualization support, and density make it suitable for general-purpose compute roles in standard SDNO use cases, while the EPAC/EUPILOT module provides targeted acceleration for ML/AI and HPC workloads in M-FLW configurations. Their compatibility with standard hardware formats and cloud software stacks allows for seamless deployment in our infrastructure without requiring architectural changes.

These systems represent a forward-looking enhancement to our cloud services. Leveraging ARM and RISC-V architectures aligns with our goals for power efficiency, sustainability, and adoption of open hardware, enabling us to deliver performant and flexible cloud infrastructure that meets evolving industry and environmental demands.

6 Conclusion and Next Steps

This document will guide the work of later WPs in which the respective parts are developed. Tasks T3.1 and T3.2 will make use of the description of the HPM processor modules, the descriptions of the DC-SCM module and Root Of Trust will be used in T3.3, and T3.4 will make use of described server mechanics. Additionally, the description of the targeted use case of CXL memory disaggregation gives further guidance for the design process of those tasks. (Respective deliverables concerning Rhea2 are due in M14 (D3.1, D3.3), those concerning EPAC/EUPILOT are due in M20 (D3.2, D3.4). The deliverables concerning the server mechanics are due in M23 and M36 (D3.5, D3.6).)

Furthermore, the description of the associated system software will inform the related tasks like T4.1 and T4.2, which will find descriptions for developing the secure boot and OS for ARM and RISC-V HPMs respectively (due for release as D4.1, D4.2 and D4.3 in M16, M22 and M28 respectively). Also, T4.3 will find guidance for the development of an open-source meta operating system that supports heterogeneous computing environments (due for release as D4.4 and D4.5 in M24 and M31 respectively).

7 Appendix

7.1 Acronyms and Abbreviations

Term	Definition
ACL	Access Control List
AI	Artificial Intelligence
AP	Application Processor
AWS	Amazon Web Services
BMC	Baseboard Management Controller
CAGR	Compound Annual Growth Rate
CCI	Common Chassis Interval
CHI	Coherent Hub Interface
CML	Coherent Mesh Link
CPM	CXL-memory Pool Manager
CRPS	Common Redundant Power Supply
CSP	Cloud Service Providers
CXL	Compute Express Link
DC-SCI	Datacenter-ready Secure Control Interface
DC-SCM	Datacenter-ready Secure Control Module
DC-MHS	data centre Modular Hardware System
DNO	DeNsity Optimized
DoA	Description of Actions
EPAC	European Processor ACcelerator
EPI	European Processor Initiative
FLW	Full Width
GCC	GNU Compiler Collection
GPP	General Purpose Processor
GPU	Graphical Processing Unit
HPC	High Performance Computing
HPE	Hewlett Packard Enterprise
HPM	Host Processor Module
ISA	Instruction Set Architecture
KVM	Keyboard-Video-Mouse
LCP	Local Control Processor
LLVM	Low-Level Virtual Machine
LTPI	LVDS Tunnelling Protocol & Interface
MCP	Manageability Control Processor
MCTP	Management Component Transport Protocol
ML	Machine Learning
MPI	Message Passing Interface
NDP	Near-Data Processing
NIC	Network Interface Card
NVMe	Non-Volatile Memory Express
OAM	OCP Accelerator Module
OCP	Open Compute Project
ODM	Original Design Manufacturer
OEM	Original Equipment Manufacturer

Term	Definition
OpenMP	Open Multi-Processing
ORv3	OpenRack v3
PCB	Printed Circuit Board
PCIe	Peripheral Component Interconnect express
PDB	Power Distribution Board
PESTI	Peripheral Sideband Tunnelling Interface
PIC	Platform Infrastructure Connectivity
PSU	Power Supply Unit
RBAC	Role-Based Access Control
RISC	Reduced Instruction Set Computer
RISE	RISC-V Software Ecosystem
RSE	Runtime Security Engine
RSS	Runtime Security Subsystem
SDNO	Scalable DeNsity Optimized
SCP	System Control Processor
SGA	Specific Grant Agreement
SME	Small and Medium-sized Enterprise
SMP	Symmetrical Multi-Processing
SNMP	Simple Network Management Protocol
TDP	Total Dissipation Power
TPM	Trusted Platform Module
TRL	Technical Readiness Level
UART	Universal Asynchronous Receiver Transmitter
UBB	Universal BaseBoard
USB	Universal Serial Bus
UEFI	Unified Extensible Firmware Interface
XIO	eXtended I/O

TABLE 3: ACRONYMS AND ABBREVIATIONS

7.2 References

Reference	Description	Version
[M-FLW]	M-FLW Base Specification	1.2RC3
[M-DNO]	M-DNO Base Specification	1.1RC2
[M-SDNO]	M-SDNO Base Specification	1.0RC2
[M-XIO]	M-XIO Base Specification	1.04RC1
[DC-MHS]	DC-MHS Specifications	NA
[OCP-NIC]	OCP NIC 3.0 Specification	1.5.0
[OAI-OAM]	OAI-OAM Base Specification r2.0	1.0
[DC-SCM]	OCP DC-SCM Specification	Rev 2.1 Version 1.1
[M-PIC]	M-PIC Base Specification	1.11
[ORV3]	Open Rack V3 Base Specification	1.0
[OCP-LTPI]	OCP LTPI Specification	1.0
[SBMR]	Arm Server Base Manageability Requirements	2.1
[SystemReady]	Arm SystemReady Requirements Specification	3.0

[D2.1]	D2.1 Requirements and use cases refinement	1.0
[CXL]	CXL Specification	3.2

TABLE 4: REFERENCES