OCP SUMMIT

March 20-21
2018
San Jose, CA

OPEN
Compute Project

OCP NIC 3.0 Collaboration

- An Open Hardware development Story

Joshua Held / Mechanical Engineer

Yueming Li / Thermal Engineer

Jia Ning / Hardware Engineering Manager

Facebook, Inc

Agenda

- Overview of project in the past one year
- OCP NIC 3.0 Mechanicals
- OCP NIC 3.0 Thermal

Background

OCP Mezzanine Cards
All accepted by the IC

Specification	Version	Submit Date	Contributor	License	Notes
OCP Mezzanine card v2.0 OCP_Mezz_2.0_rev1.00_20151215b_pub_release.pdf(2.2MB) OCP_Mezz_2.0_rev1.00_20151215b_pub_release_3D_package.zip (88MB) Mechanical 20151023_P1-P9_K1-K5 zip file (57MB)	V2.0-1.0	Dec 15, 2015	Facebook	OWFa 1.0	Added support for x16 (quad x4), NCSI, dual QSFP+, & Quad SDP+ Accepted by OCP IC 2/24/2016
OCP Mezzanine card v0.5, original standard Mezzanine Card (rev 0.5)	V0.5	Oct 8, 2012	Facebook	OWFa 1.0	Defacto standard for the original network mezzanine with a x8 PCIe Gen3 interface

OCP Mezz v0.5

- Defined ~2012
- 10G Ethernet
- 2x SFP
- X8 PCIe Gen3
- I2C sideband

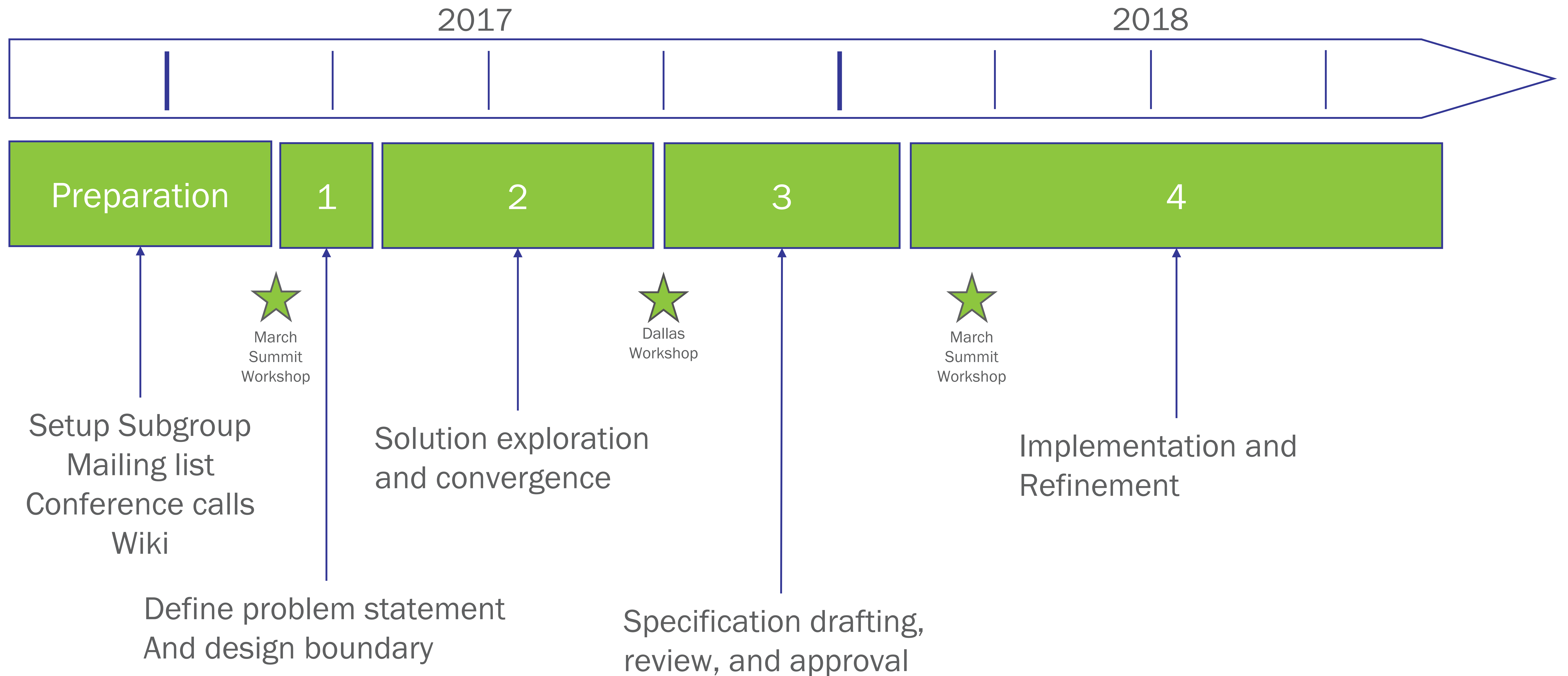
OCP Mezz v2.0

- Defined ~2015
- 10/25/40/50/100G Ethernet
- Up to 4x SFP28, 2x QSFP28, 4x RJ45
- X16 PCIe Gen3
- NCSI sideband

OCP NIC 2.0 Limitation

- Gates emerging use cases & blocks broader adoption
  - Board space
  - Mechanical and thermal profile
  - Connector placement
  - Specification quality

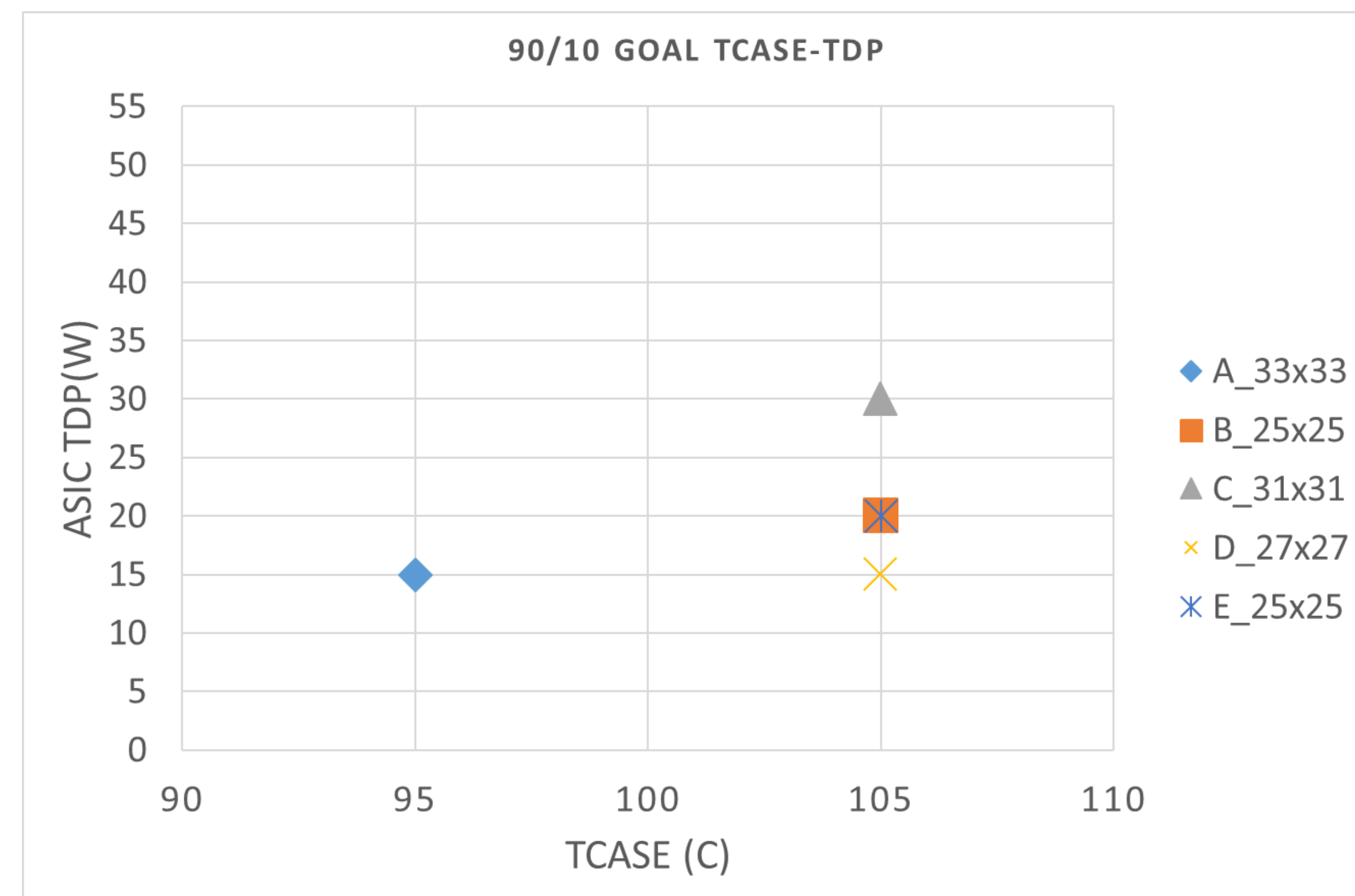
OCP NIC 3.0 Milestones

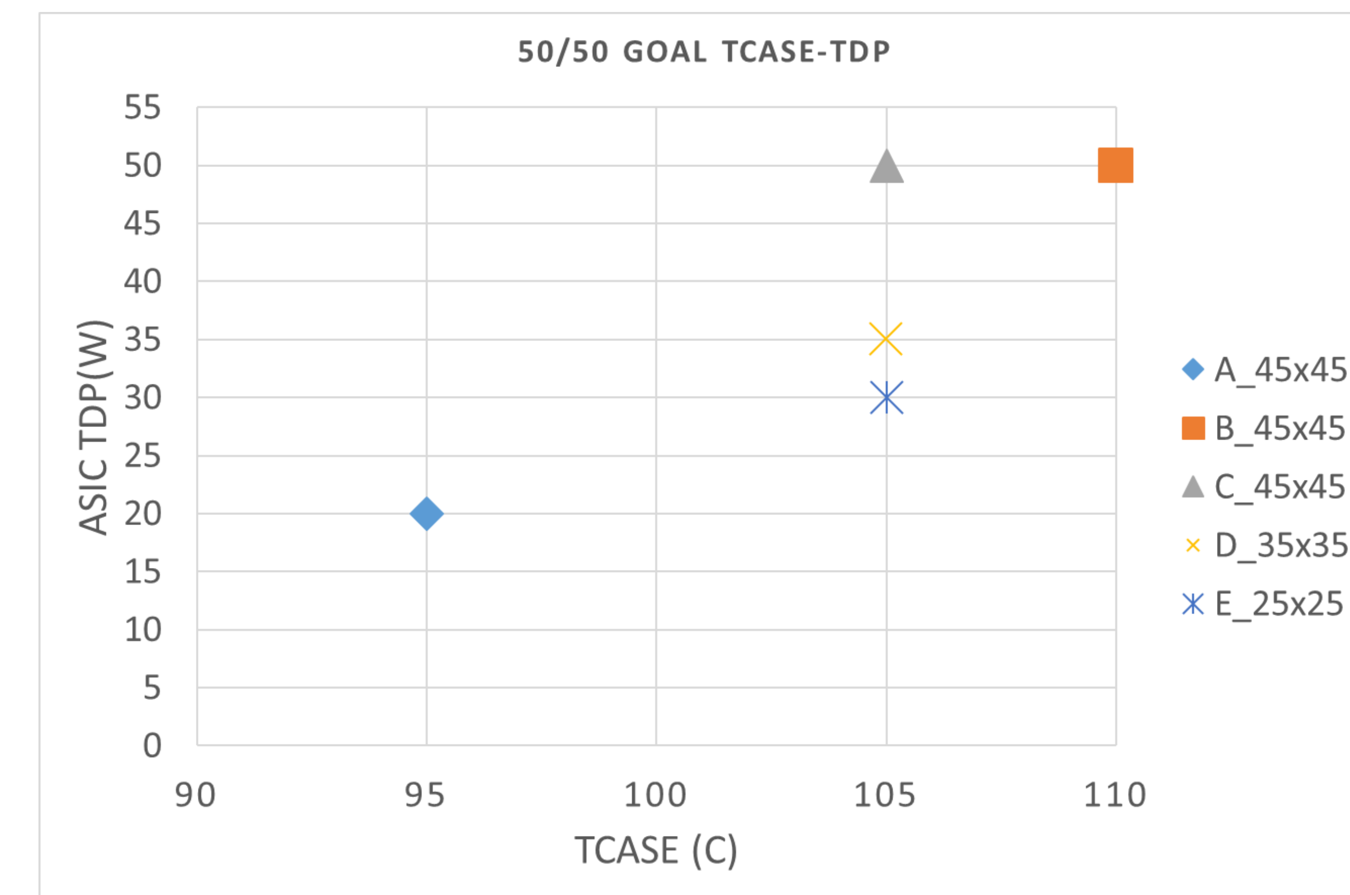


Define Problem Statement

March to Mid-April '17

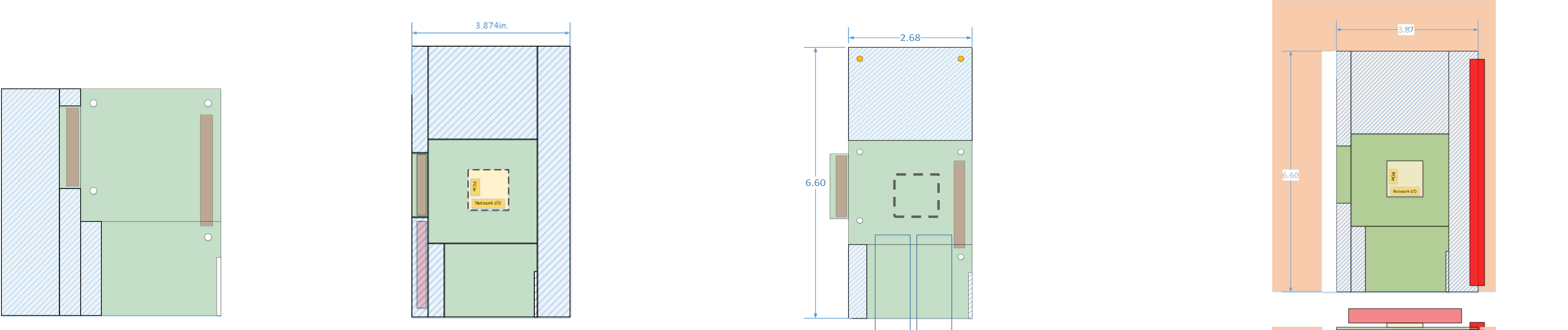
OCP Partner	A		B		C		D		E	
	90/10 Very typical and Important use cases	50/50 confidences Stretch goals and use cases	90/10 Very typical and Important use cases	50/50 confidences Stretch goals and use cases	90/10 Very typical and Important use cases	50/50 confidences Stretch goals and use cases	90/10 Very typical and Important use cases	50/50 confidences Stretch goals and use cases	90/10 Very typical and Important use cases	50/50 confidences Stretch goals and use cases
# and Type of **I/O** Ports	2, (2xSFP/SFP28, 2xQSFP/QSFP28)	2, (2xSFP/SFP28, 2xQSFP/QSFP28)	1-4 ports SFP+ or Base-T, or 1-2 QSFP	1-4 ports SFP+ or Base-T, or 1-2 QSFP	Dual port(side by side/Belly to Belly) 200Gb	Dual port(side by side/Belly to Belly) 200Gb	2 x SFP		4 *SFP , or 2*QSFP, or 4*FC, or 2*SAS	
# of major IC (ASIC, FPGA, or other)	1	1	1 or 2 (ASIC, ASIC+PHY)	2 or 3 (ASIC + FPGA/SoC + PHY)	1	2	2	2	2 (controller + Phy)	
# of DRAM	6 x16	12 x16	10	20	N/A	N/A	0 or 2	2+	0 or 2	
Power envelope of **IC#1** **(Max power at Tj max)**	15	20	20	50	30	50	15	35	20	30
Max T_case of **IC#1**	95	95	105	110	105	105	105		105	105
IC#1 mechanical dimension (WxLxH)	33x33x2.5	45x45x3.7	25x25x 3.7	45x45x3.7	31x31	45x45	27 x 27mm + IC2 (or 47 x 47mm single chip)	35 x 35mm + IC2 (or 47 x 47mm single chip)	25x25	
If Applicable:	N/A	N/A	Phy	Phy	N/A	N/A				
Power envelope of **IC#2** **(Max power at Tj max)**	N/A	N/A	3W per port	3W per port, total <50W	N/A	N/A	20W		10w	
Max T_case of **IC#2**	N/A	N/A	105	110	N/A	N/A	110		110	
IC#2 mechanical dimension (WxLxH)	N/A	N/A	19x19x~2.5 mm	19x19~2.5 mm	N/A	N/A	25 x 25		25x25	
If Applicable:	N/A	N/A	N/A	FPGA/SoC	N/A	N/A	N/A	N/A	N/A	N/A
Power envelope of **IC#2** **(Max power at Tj max)**	N/A	N/A	N/A	TBD, total < 50W	N/A	N/A	N/A	N/A	N/A	N/A
Max T_case of **IC#2**	N/A	N/A	N/A	105	N/A	N/A	N/A	N/A	N/A	N/A
IC#2 mechanical dimension (WxLxH)	N/A	N/A	N/A	25x25x3.7	N/A	N/A	N/A	N/A	N/A	N/A
If Applicable:										
DRAM Power (each component **at Tj max**)	0.33	0.5	0.4	0.4						
Max T_case of DRAM	95	95	95	95			95C		80	80
If Applicable:										
# of **Optical modules**	0	2	1-4 SFP+, or 1-2 QSFP	1-4 SFP+, or 1-2 QSFP			4		4	
Optical Module power (each)	0	1.5	1.5 watts (SFP+), 3.5 watts (QSFP)	1.5 watts (SFP+), 3.5 watts (QSFP)			1.5W each		1.5W each	
Optical Module **Max** Tcase	0	70	70 deg C	85 deg C			85C	70C	85C	
System air flow direction	Inlet	Inlet	Both	Both	Port to ASIC, 35C ambient	Port to ASIC, 35C ambient	Hot aisle operation; Air inlet to card max temp 55C	Hot aisle operation; Air inlet to card max temp 65C	front to rear , IO on rear, temp hitting IO up to 70C	



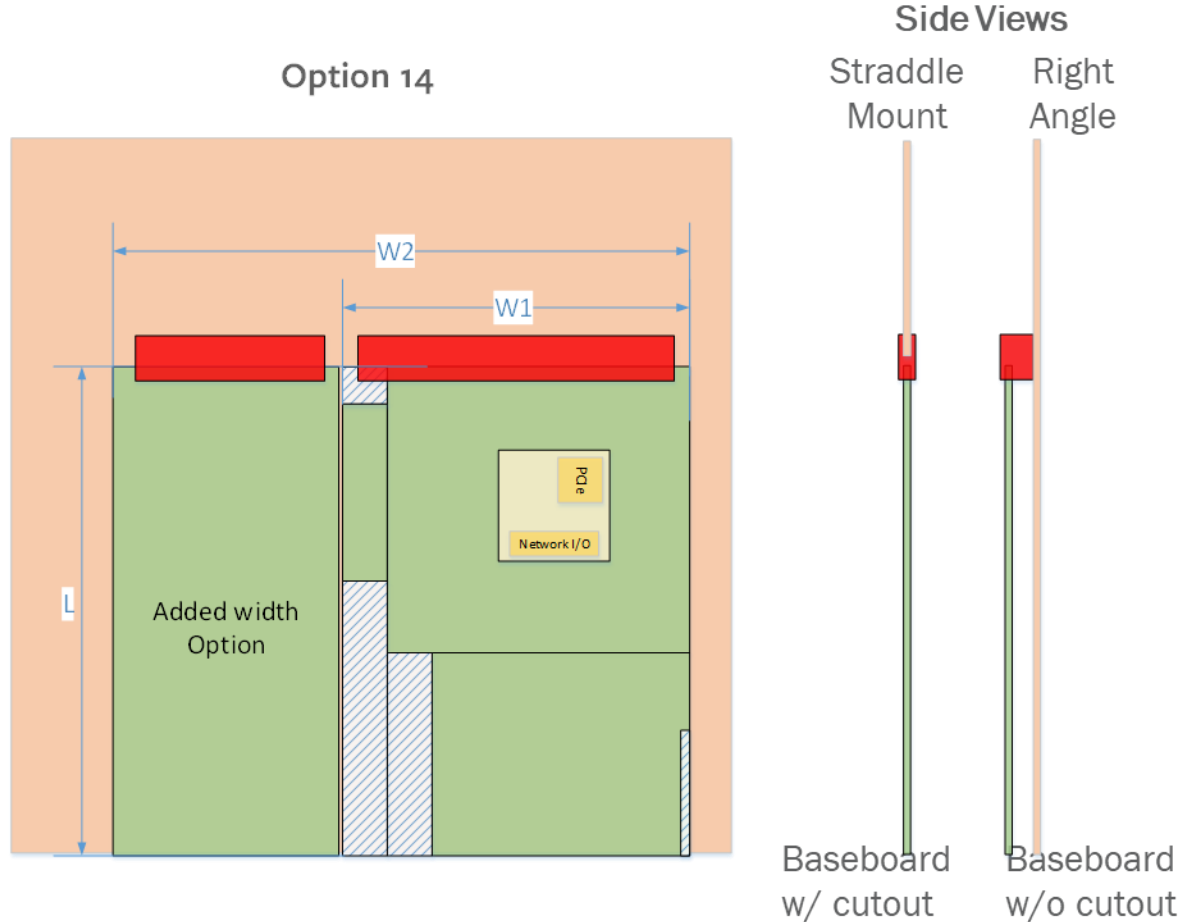


Solution Exploration and Convergence

Mid-April to end-of-September'17



14x solution options proposed and evaluated

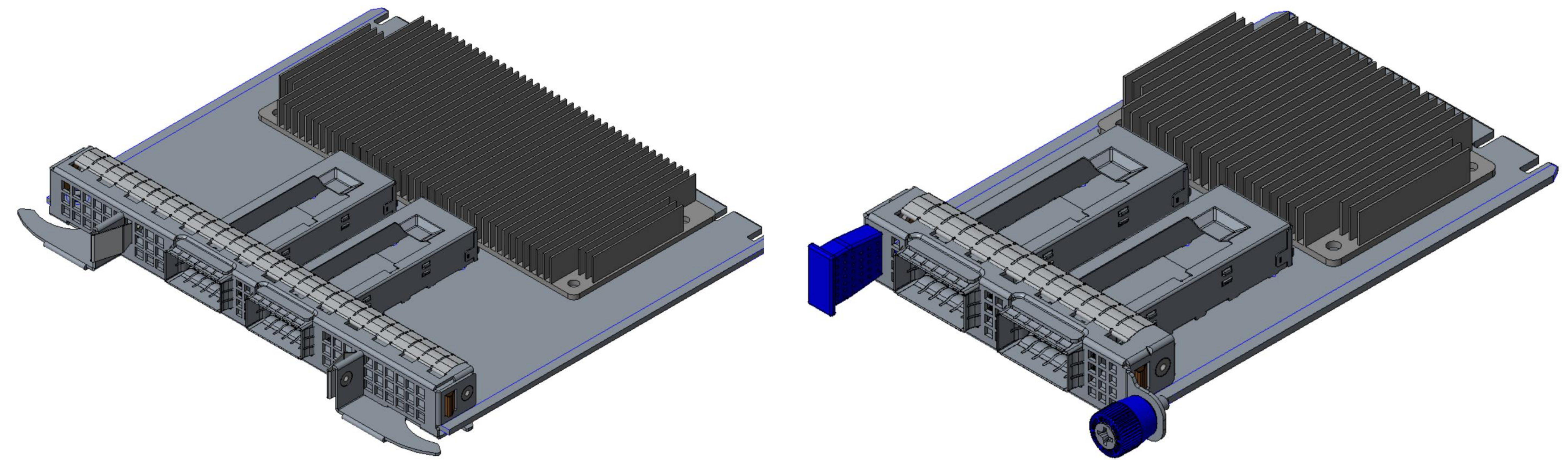


Dallas Workshop

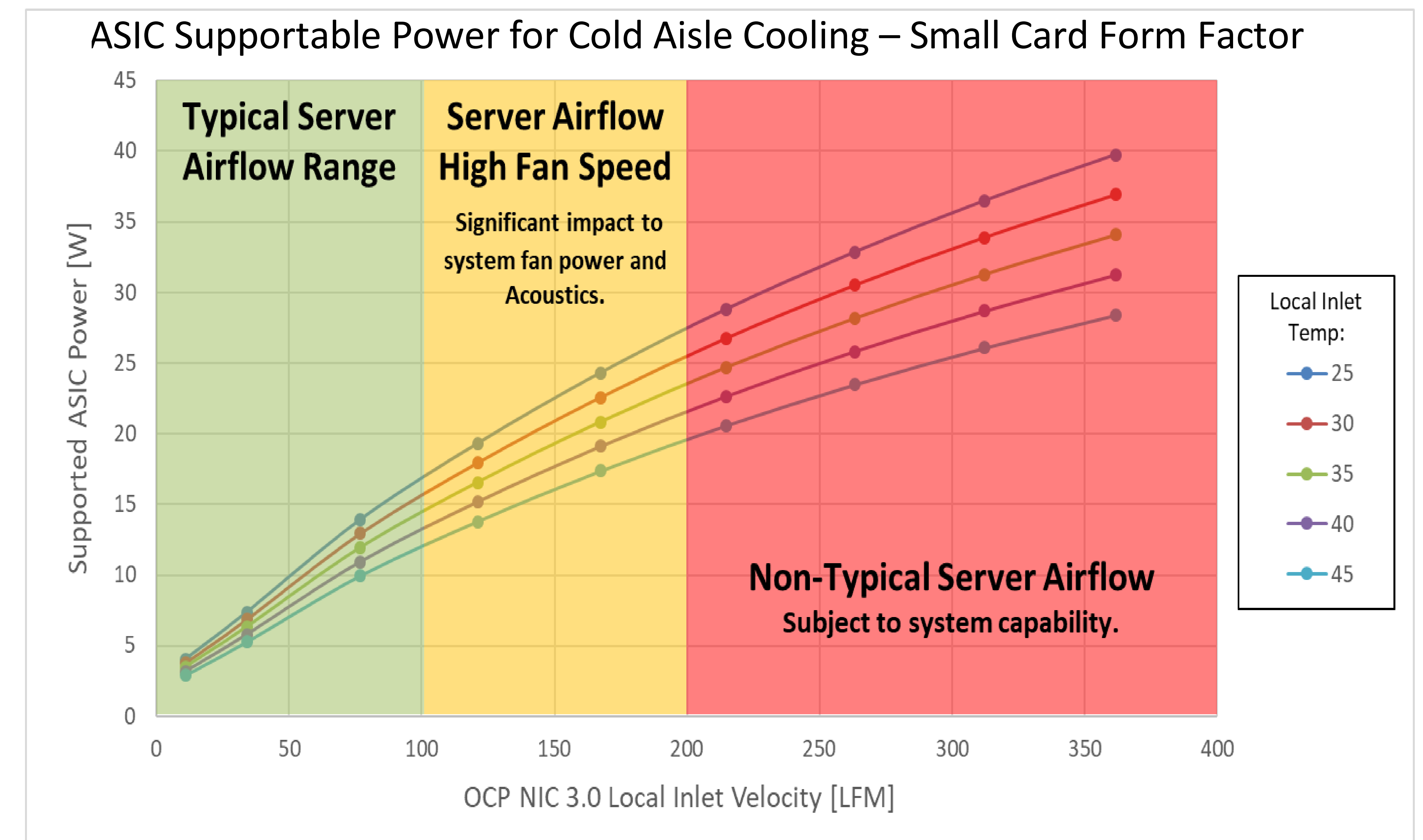
Sep 25th, 2017

- First F2F workshop
  - Met each other in person
  - Had an open discussion
  - Accelerated the collaboration

Solution Overview



- 2x Form factors (SFF and LFF)
- SFF-TA-1002 connector
- 32 lanes of PCIe Gen4
  - 4x of OCP NIC 2.0
- EMI containment
- Front service
- 80W/150W power delivery
- Larger thermal potential in similar profile
- NIC management features



Latest specification : http://www.opencompute.org/wiki/Server/Mezz

Draft – Review - Approval

Oct’17 to Jan’18

15x General specification working sessions
11x Mechanical specific working sessions
6x Thermal specific working sessions
41x doc revisions

1 Specification
Mechanical 3D models
Mechanical 2D models
Thermal simulation models
Thermal test fixture model

Community partners

Amphenol Corporation
Broadcom Limited
Dell, Inc.
Facebook, Inc.
Hewlett Packard Enterprise Company
Intel Corporation
Lenovo Group Ltd
Mellanox Technologies, Ltd
Netronome Systems, Inc.
Quanta Computer Inc.
TE Connectivity Corporation
And many more!

OPEN
Compute Project

OCP NIC 3.0 Design Specification

Version 0.70

Author: OCP Server Workgroup, OCP NIC subgroup

Implementation and Refinement

0v70 – Initial release
Jan 25th,2018

0v80 – Hot fixes
End of March'18

0v90 – Add Signal Integrity Guideline and Conformance
TBD

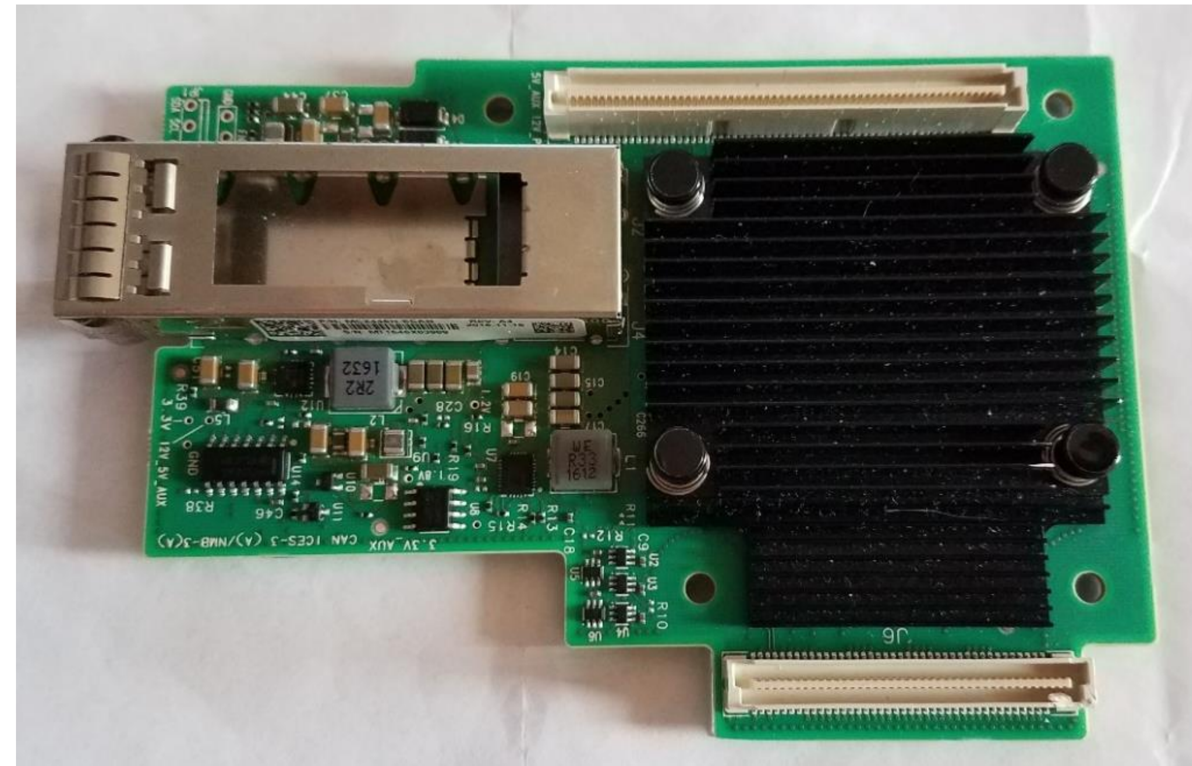
1v00 – Add Implementation Learning
Q4'18

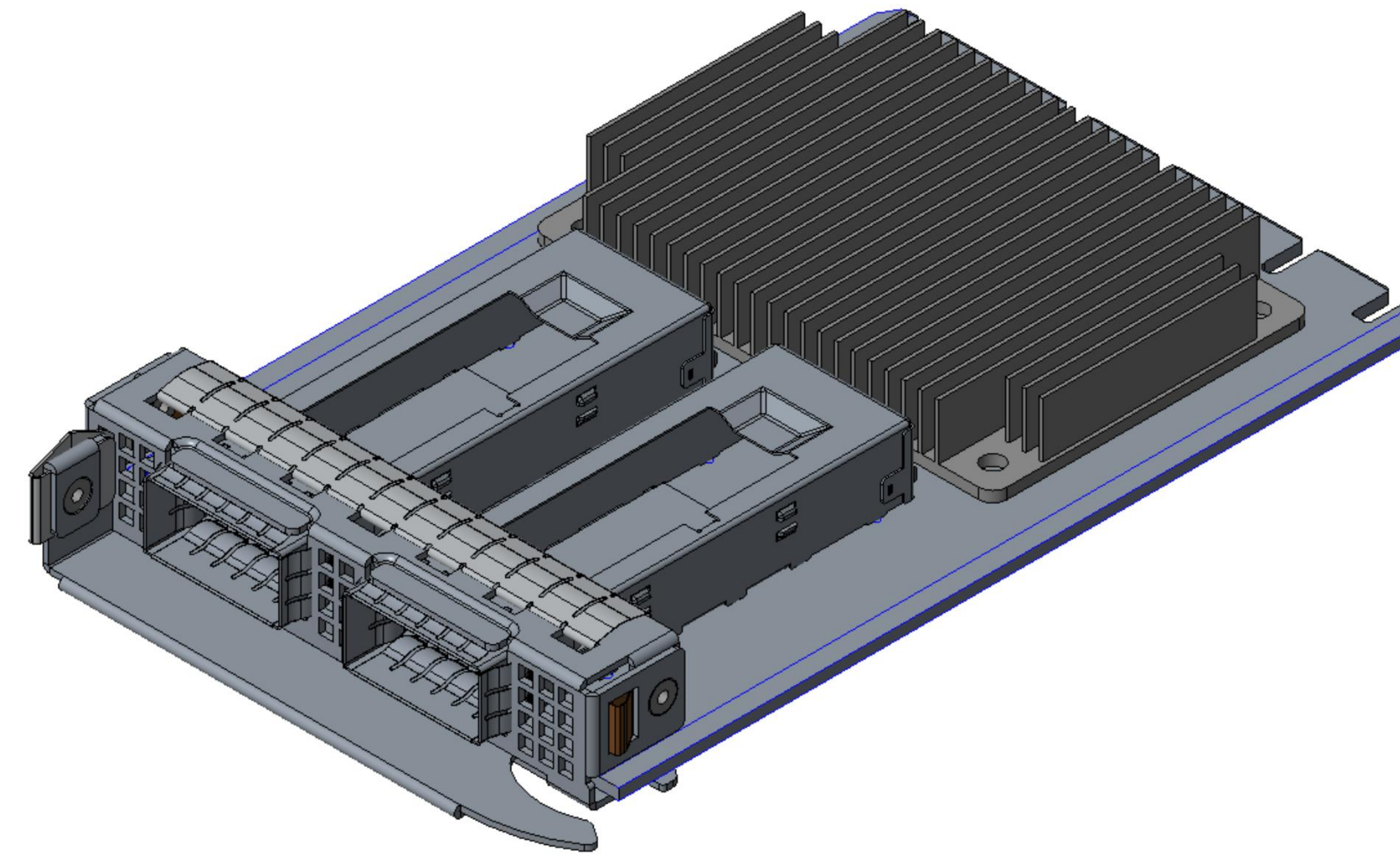
Subgroup Wiki with latest specification : http://www.opencompute.org/wiki/Server/Mezz

Mailing list: http://lists.opencompute.org/mailman/listinfo/opencompute-mezz-card

OPEN Compute Project
OCP SUMMIT

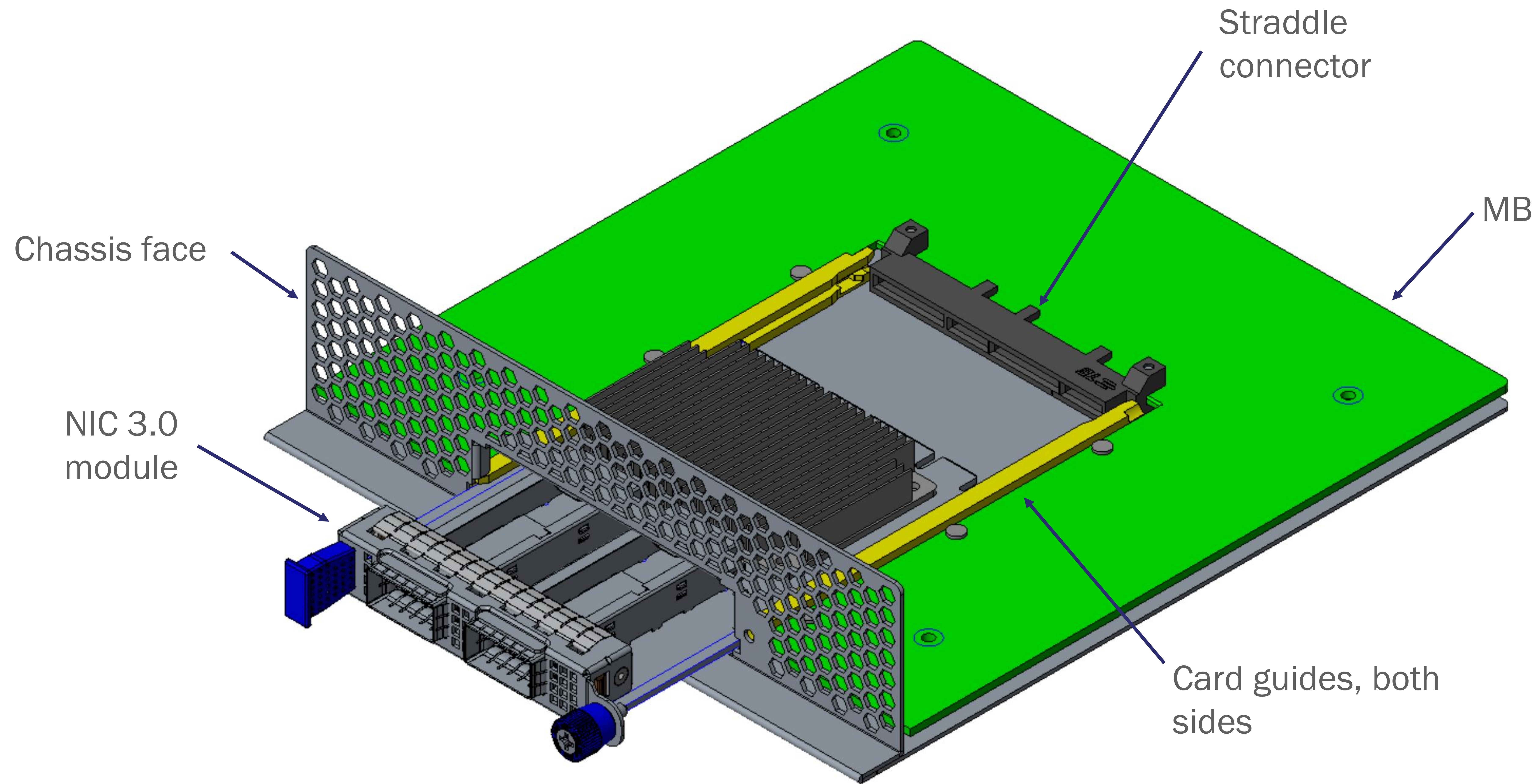
OCP Mezz 2.0 vs OCP NIC 3.0





	Mezz 2.0	NIC 3.0
Small Size	Non-Rectangle	76x115
Small Area	8000 mm^2	8740 mm^2
Large Size	NA	139x115
Large Area	NA	15985 mm^2
Expansion Direction	NA	Side
Connector style	Mezz	Edge (.6mm pitch)
PCB Orientations	Parallel	Parallel
Installation	In Chassis	Front/Rear Panel
Installation Action	Parallel to Front/Rear Panel	Perpendicular to Front/Rear Panel
Hot Swap	No	Yes
EMI Containment for Serviceability	High Difficulty	Low Difficulty

NIC 3.0 Configuration



NIC 3.0 Mechanical Goal:

Develop universal form factors which shall include mechanicals and EMI containment.

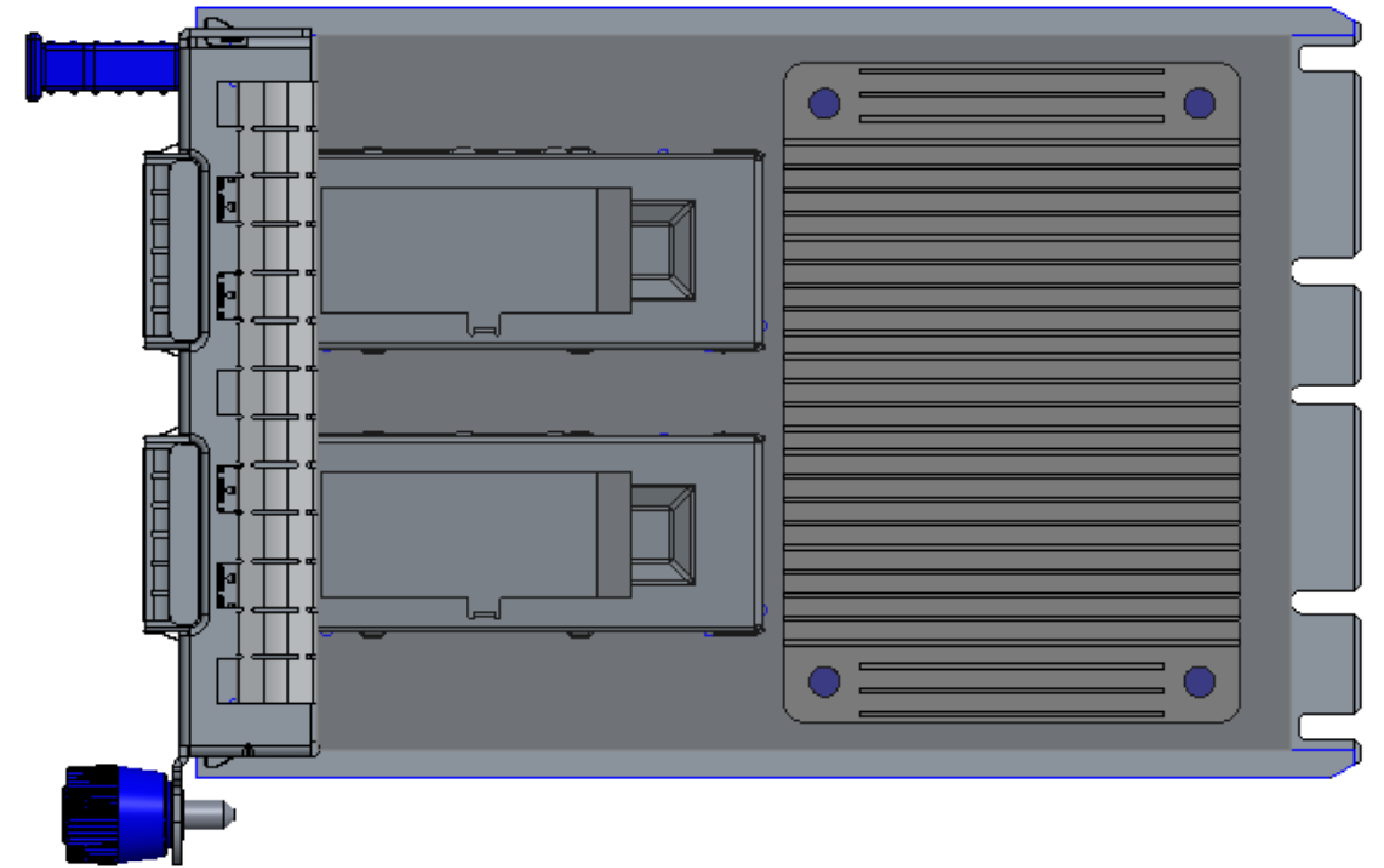
Requirements:

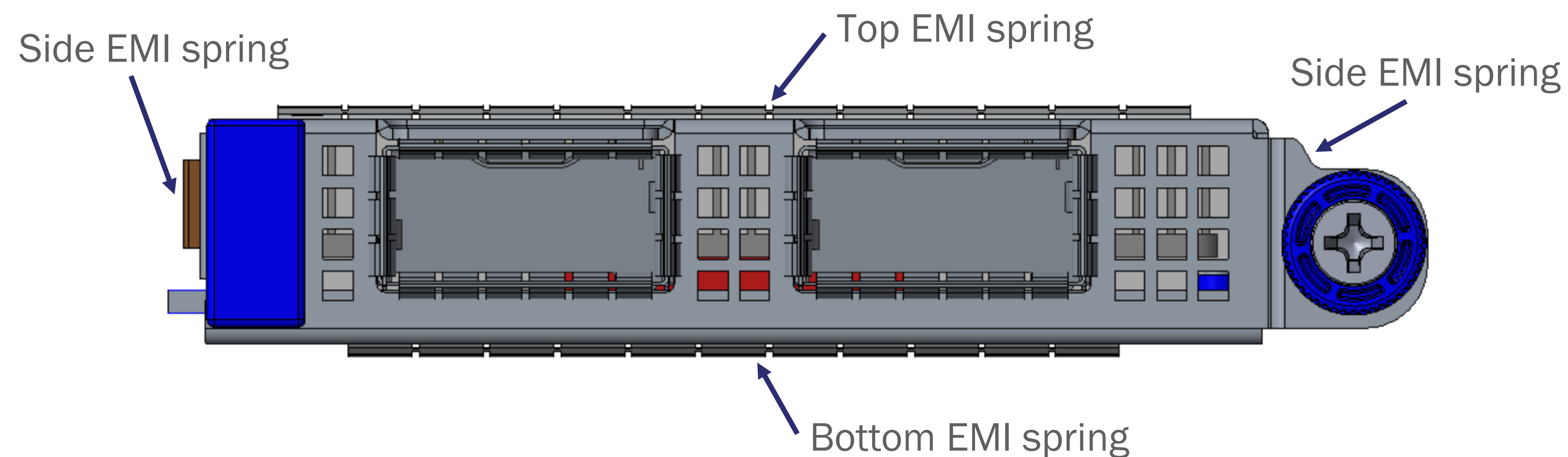
- The same NIC design must work in both straddle and right angle configurations
- NICs might be oriented horizontally or vertically
- Retention and guidance must work with specified NIC v3.0 MB thicknesses
- Common mechanical features used across small and large form factors
- Some form of mechanism is required for seating large form factors due to mating forces
- Design of PCB should be flexible enough to support many component and connector configurations without need of mechanical changes
- Recommended mechanical designs will be included in the specification to simplify and reduce barriers to adoption

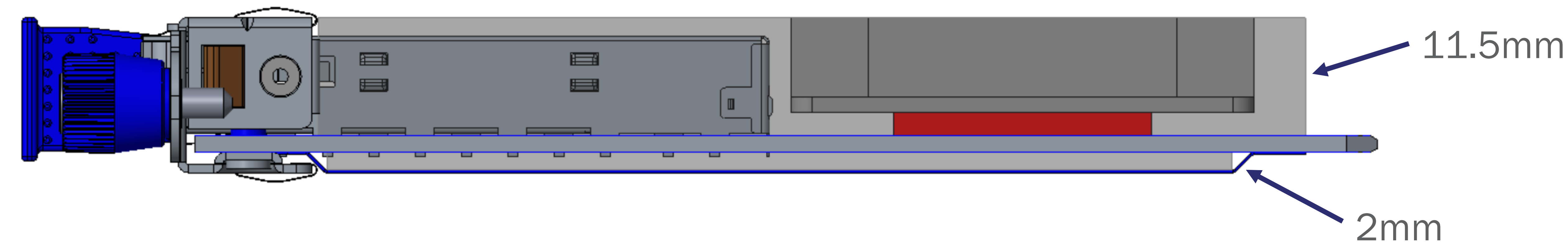
NIC 3.0 Module Configuration

Features:

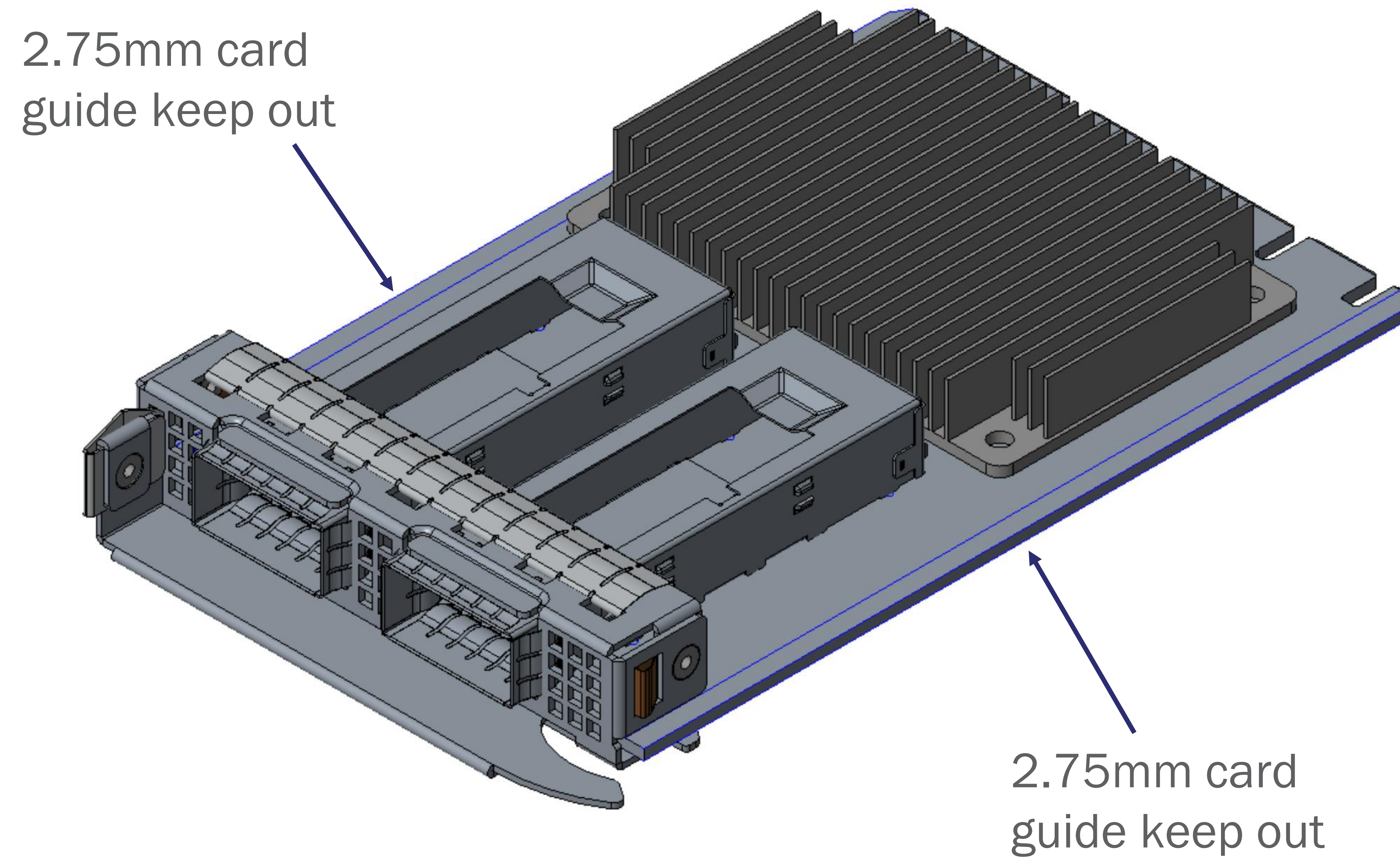
- Increased total PCB space
- Simplified component keep-in areas
- Scalable design to support large form factor
- Built in EMI containment
- Available in thumb screw or tool-less configurations with no PCB changes







NIC 3.0 Module Sizes



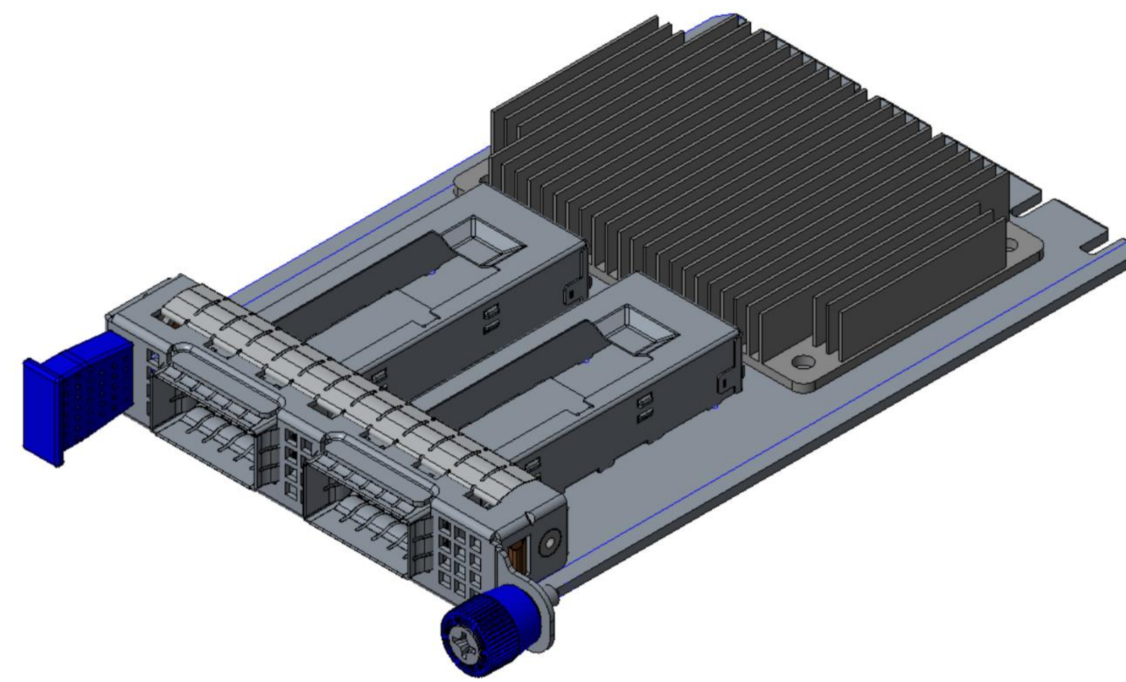
		Available Space	
	PCB Size	Top/Bottom Component Placement	Routing Inner Layers
Small Size (SFF)	76x115	70.5x109	76x115
Large Size (LFF)	139x115	133.5x109	139x115

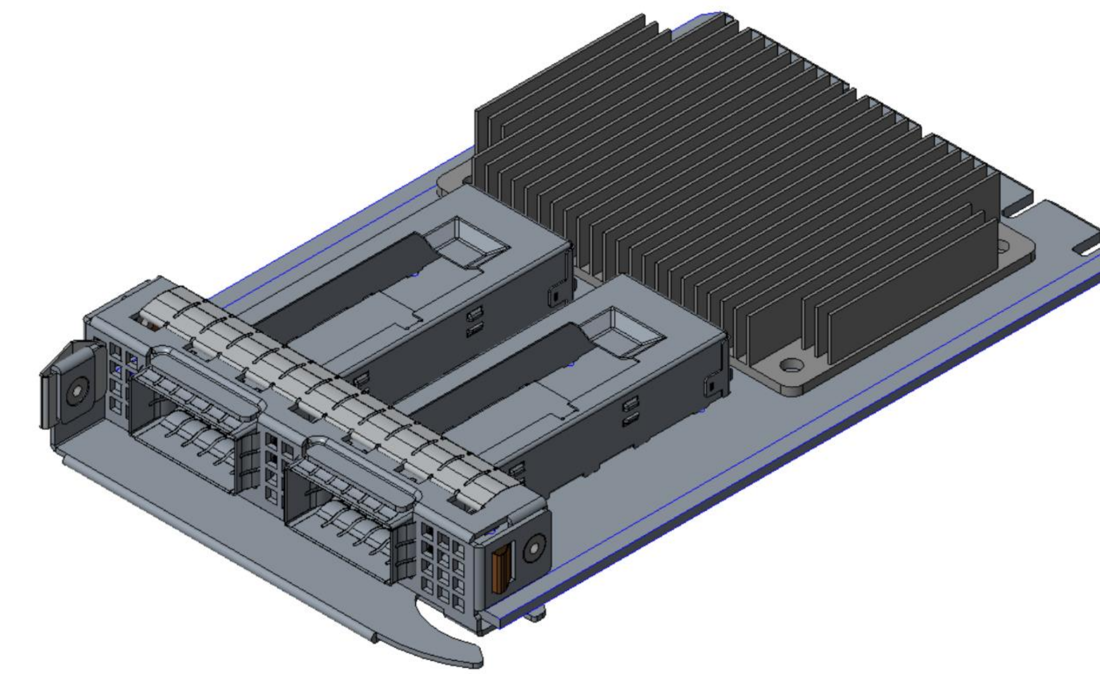
subtract 5.5mm from width (card guide)
subtract 6mm from length (edge connector)

NIC 3.0 SFF Module Versions

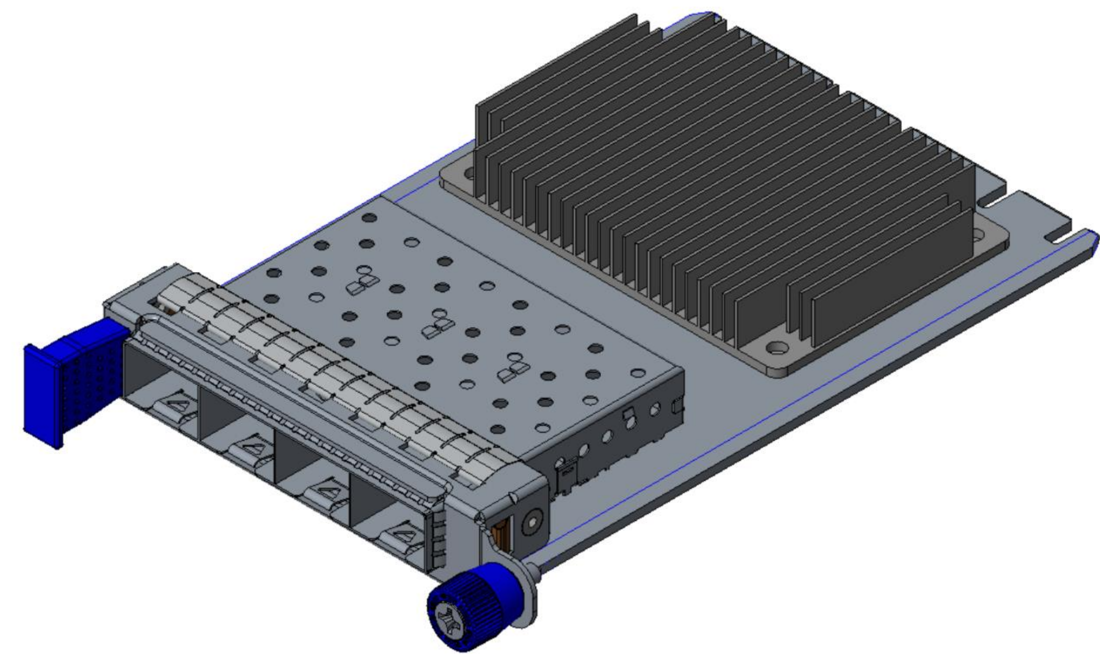
Complete 3D CAD available at:
http://www.opencompute.org/wiki/Server/Mezz

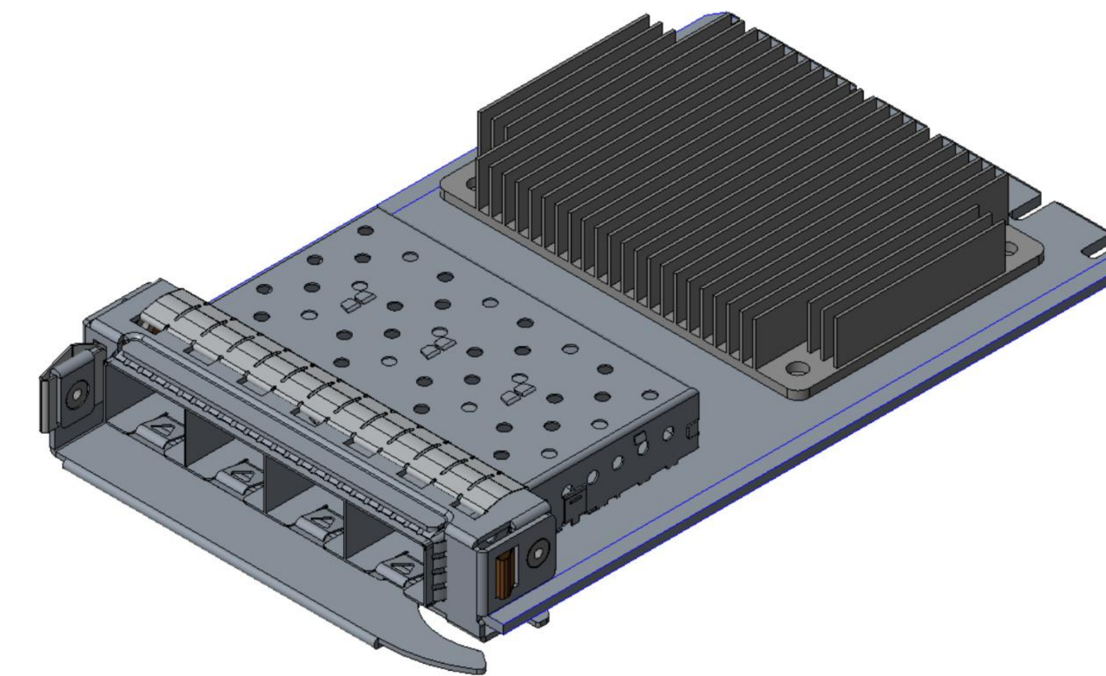
2x QSFP



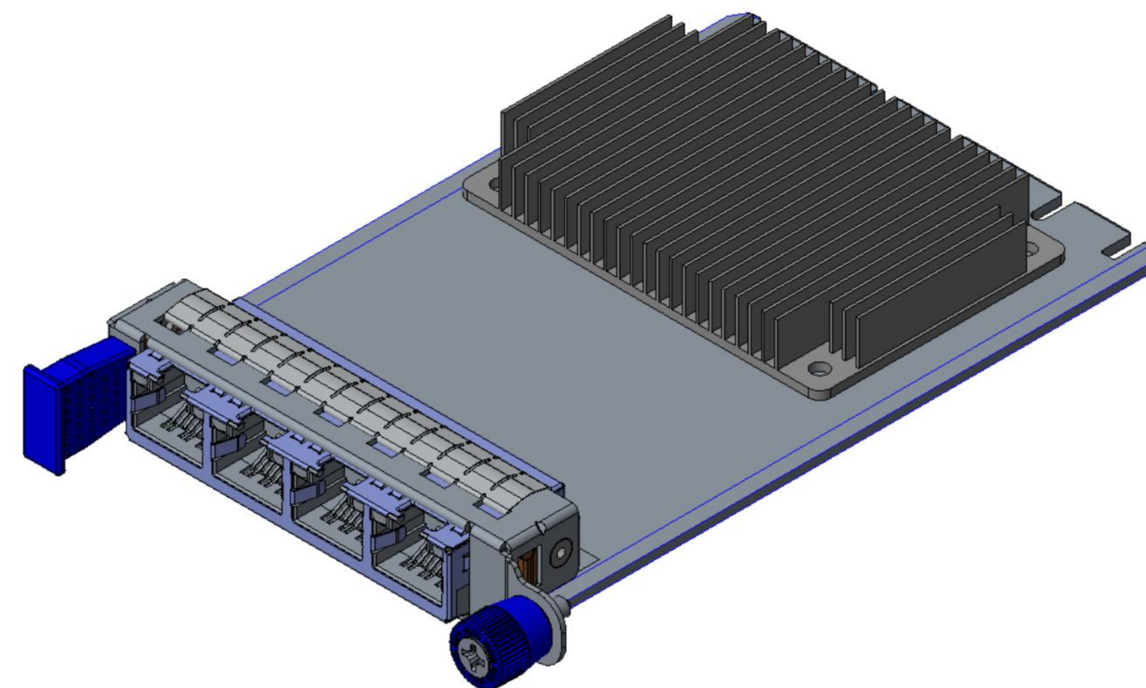


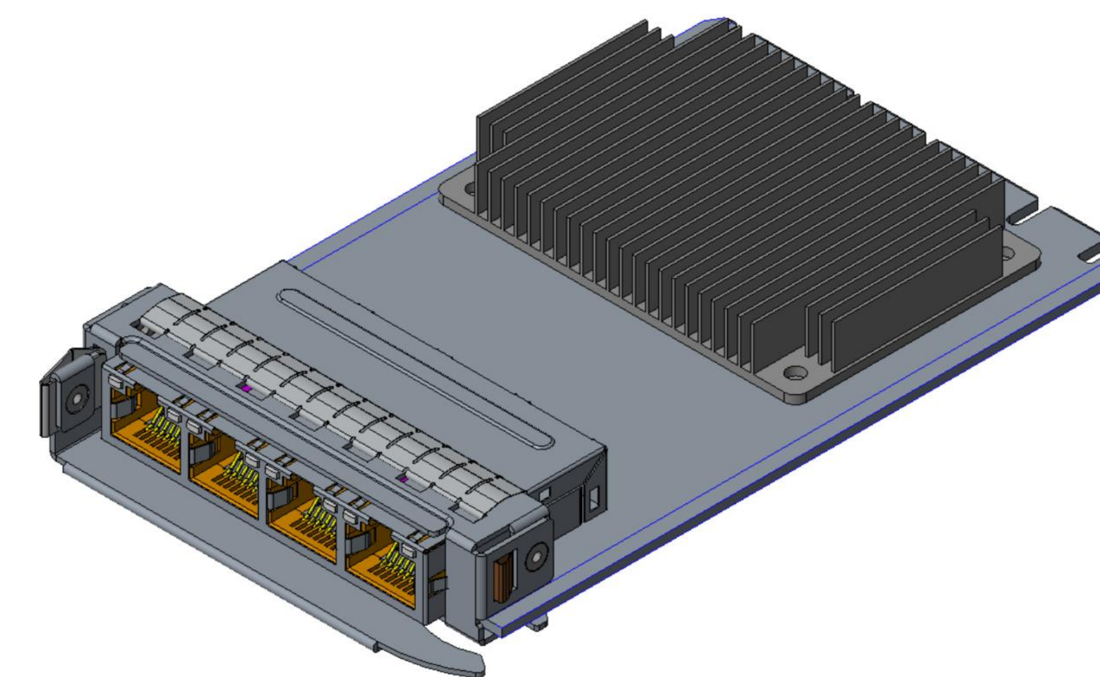
4x SFP





4x RJ45





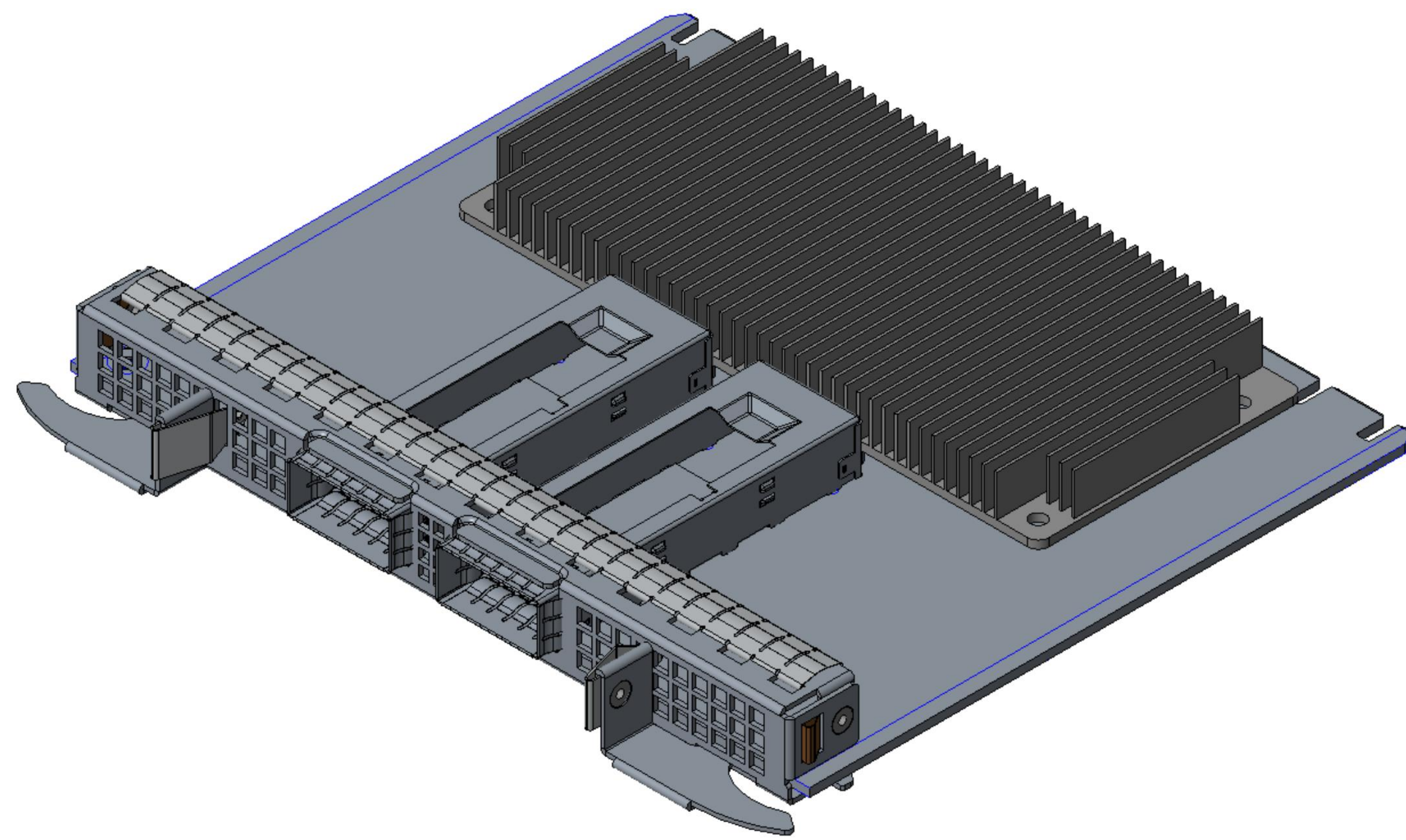
Single thumb screw version

Tool-less version

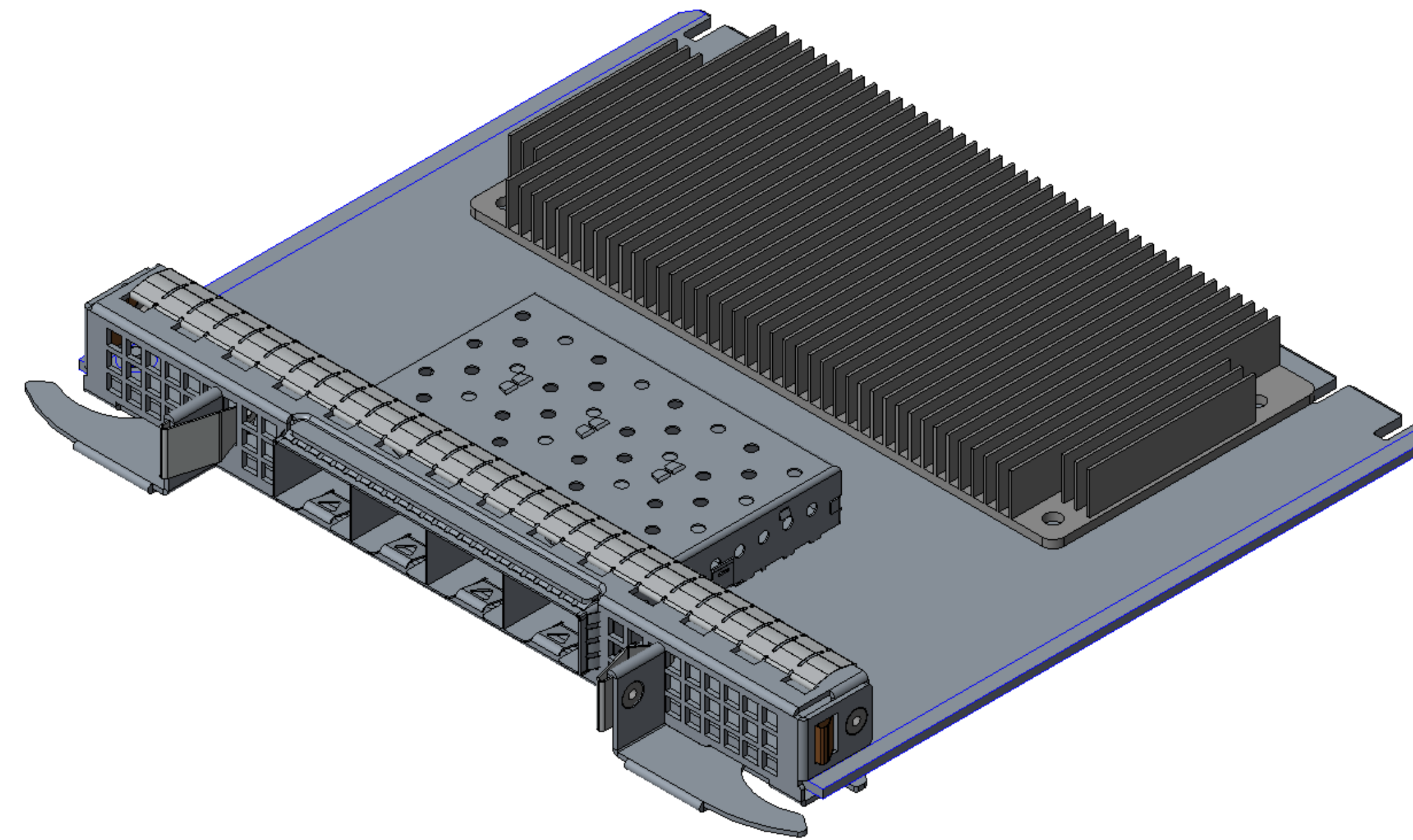
NIC 3.0 LFF Module Versions

Complete 3D CAD available at:
http://www.opencompute.org/wiki/Server/Mezz

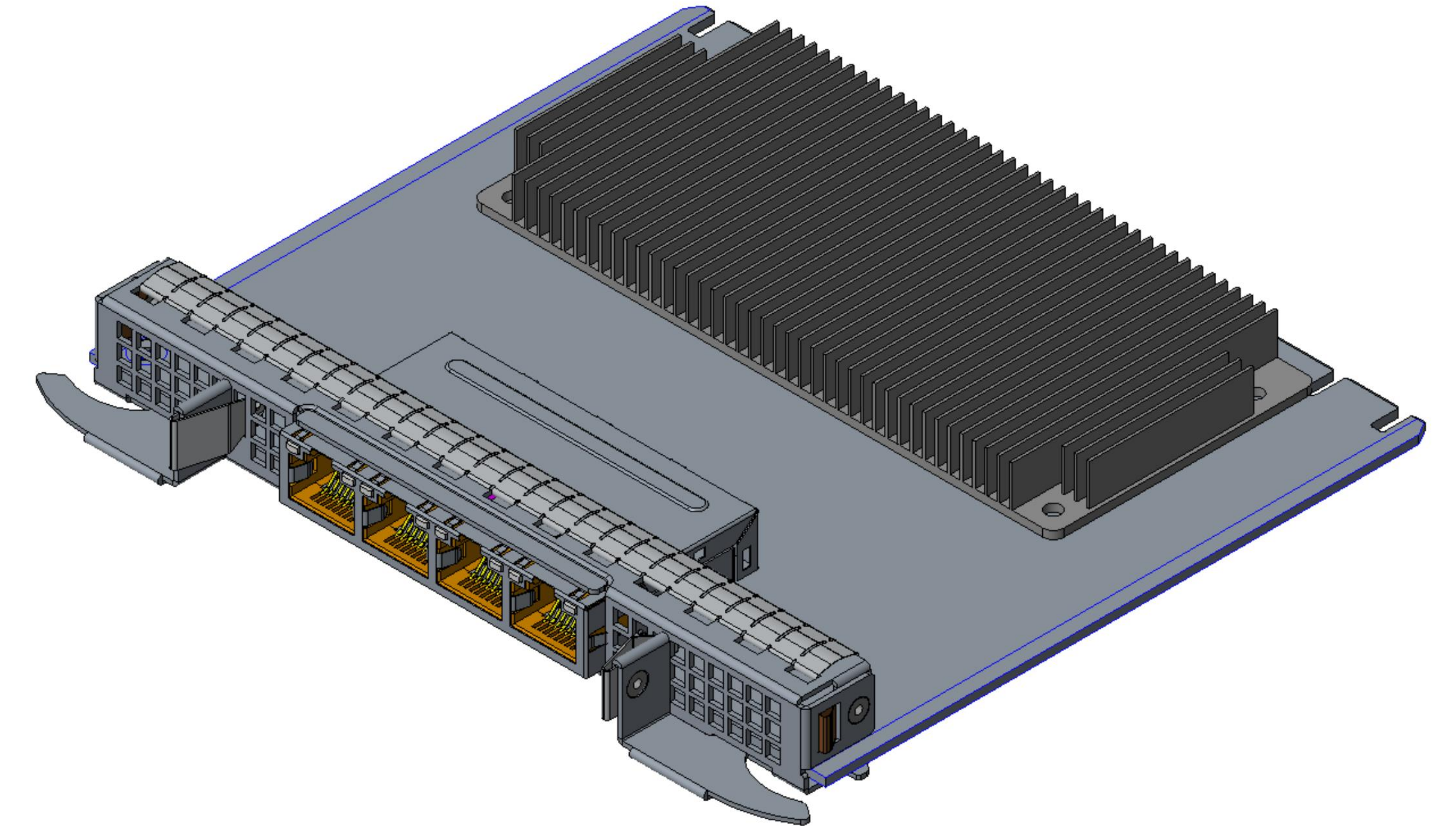
2x QSFP



4x SFP



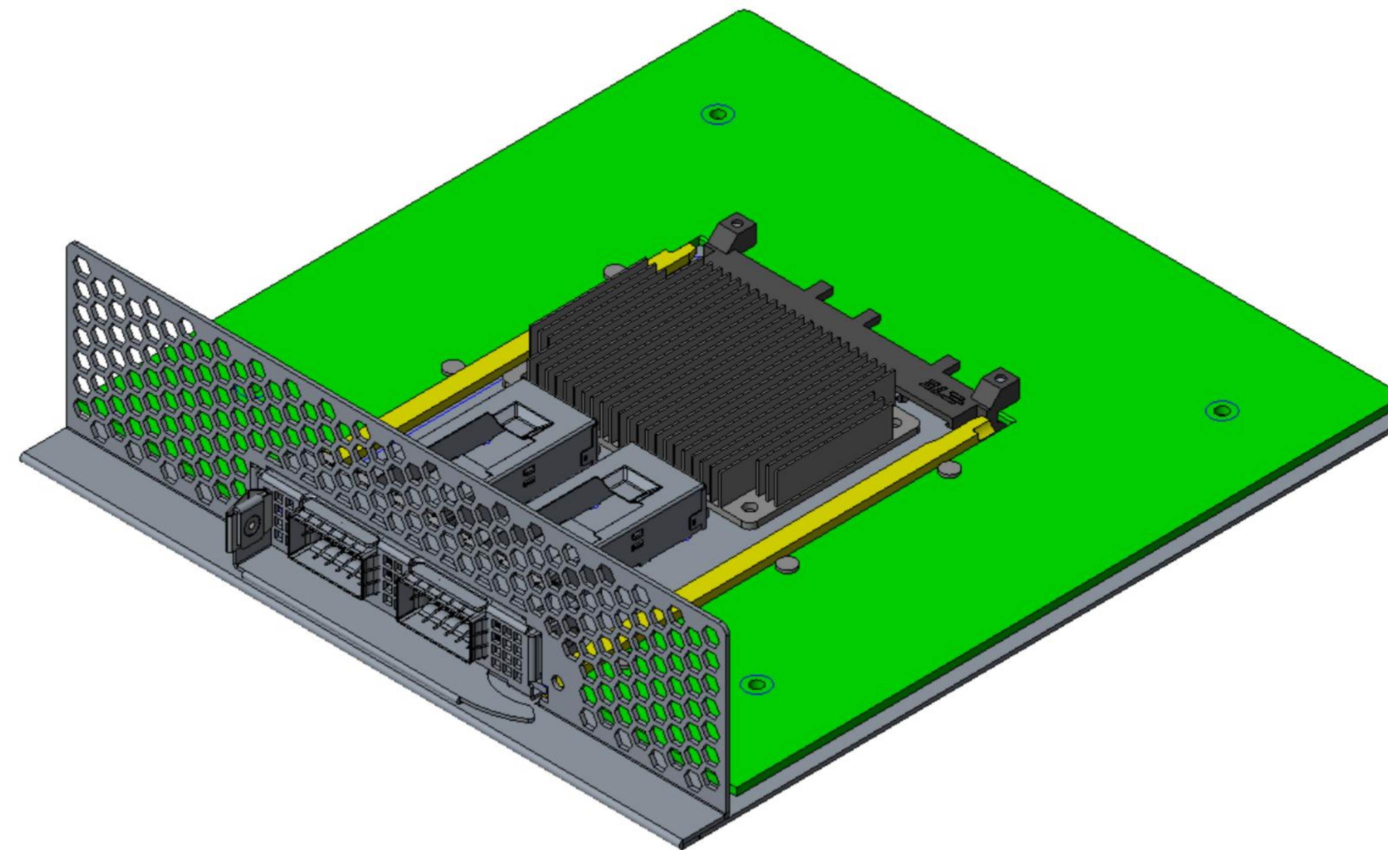
4x RJ45

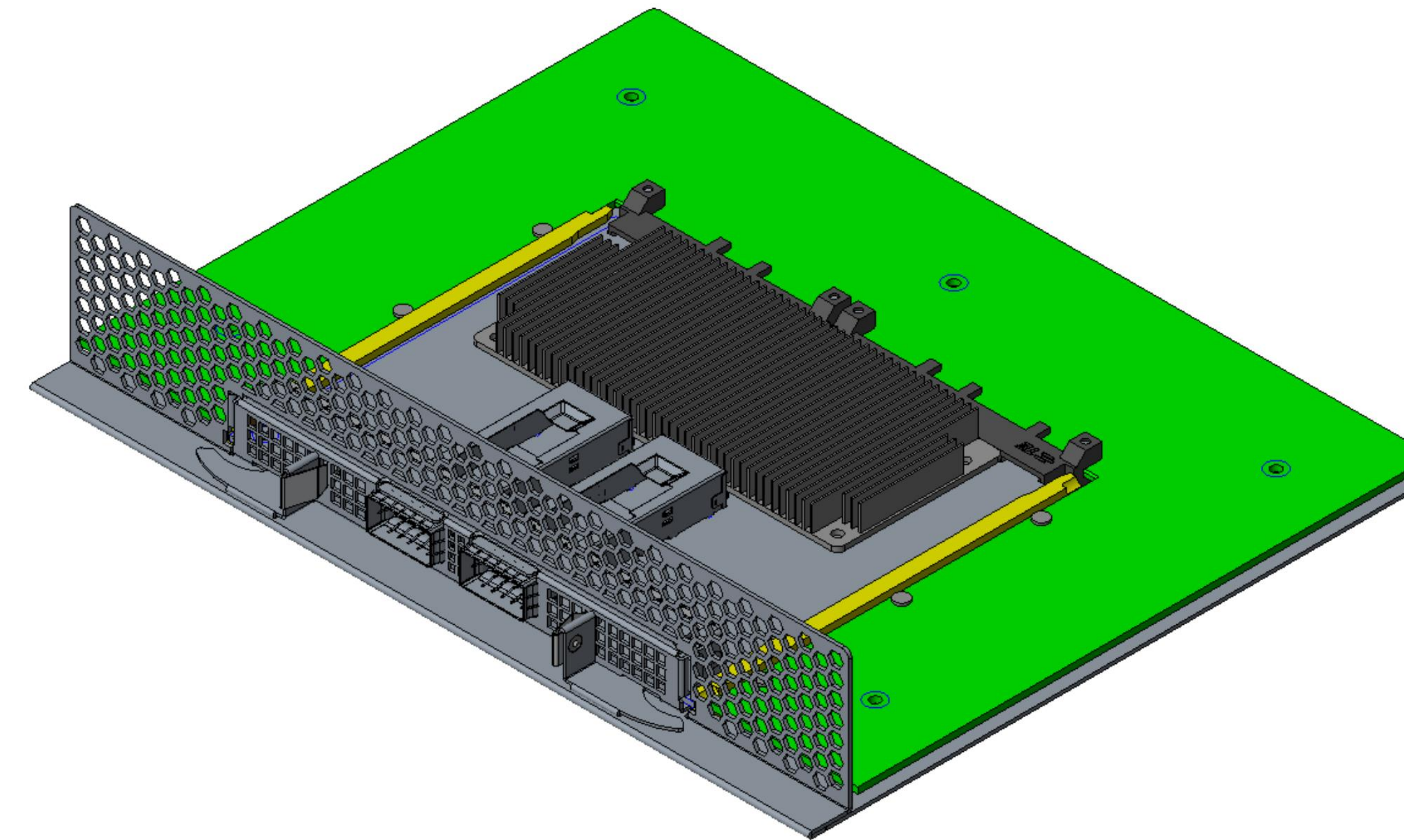


NIC 3.0 Chassis Examples

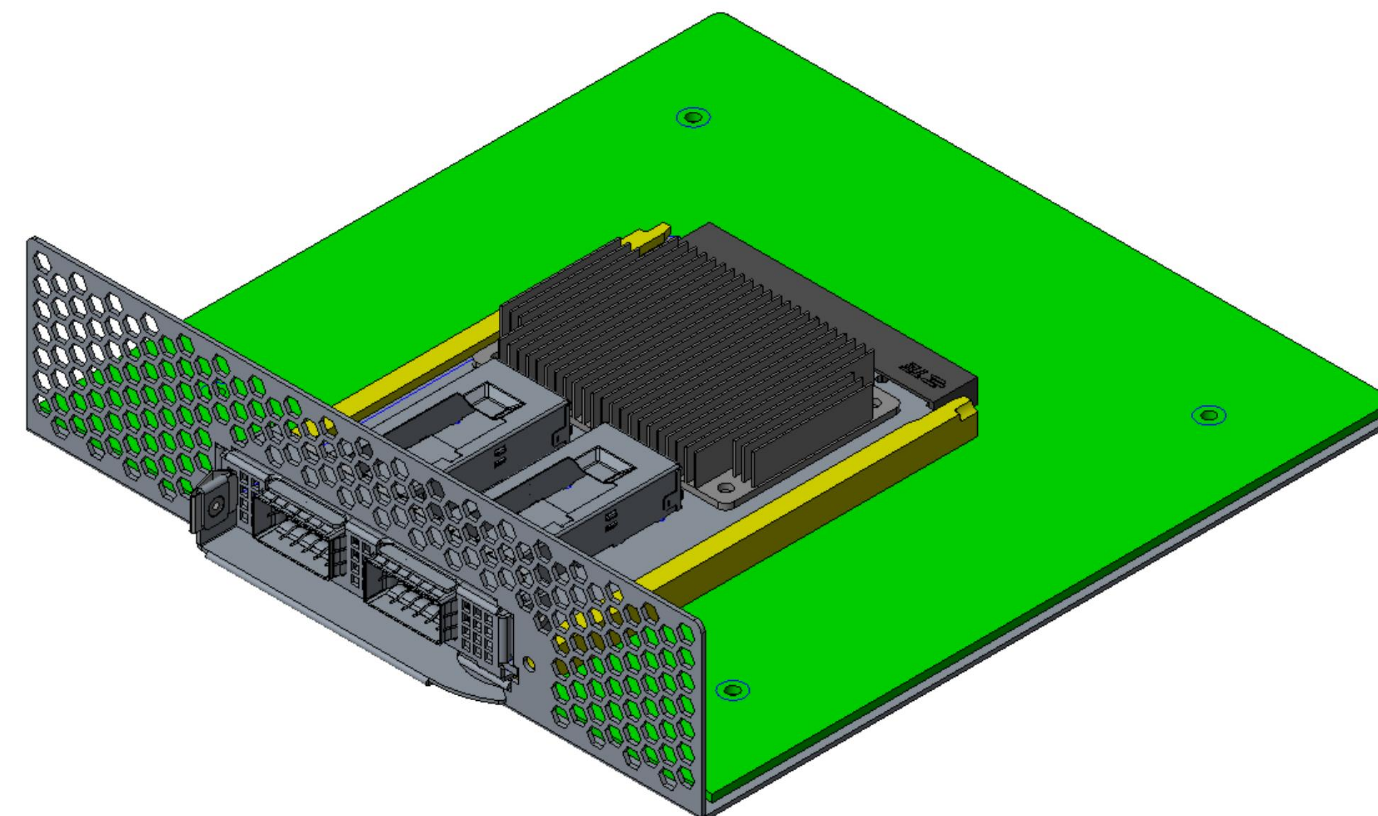
Complete 3D CAD available at:
http://www.opencompute.org/wiki/Server/Mezz

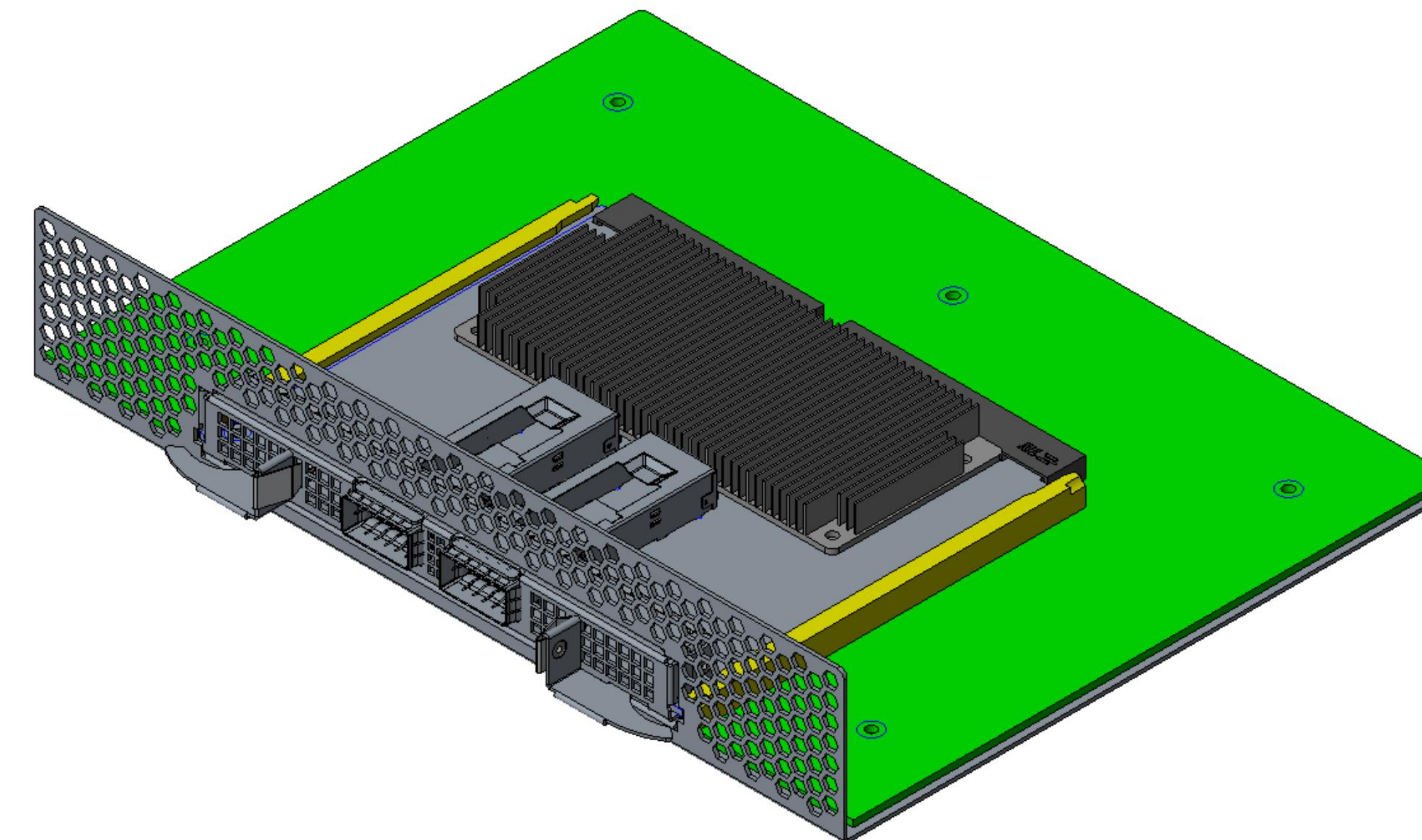
Straddle Mount





Right Angle Mount

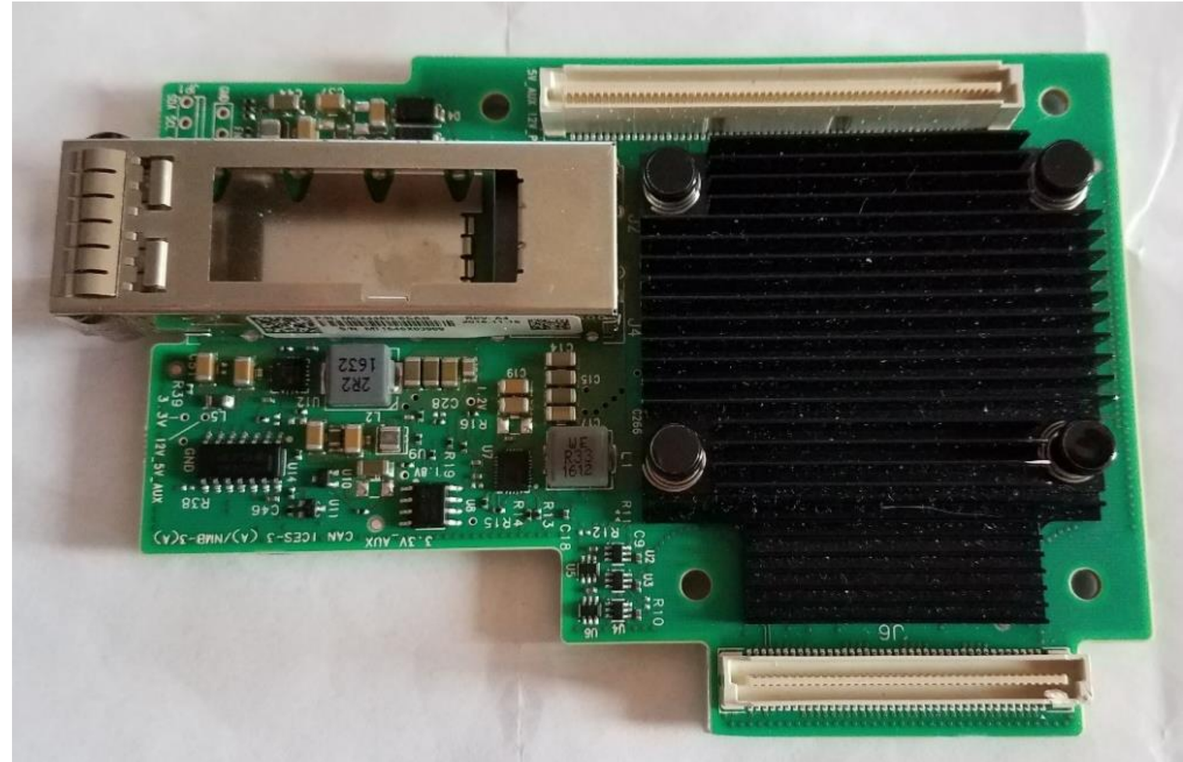




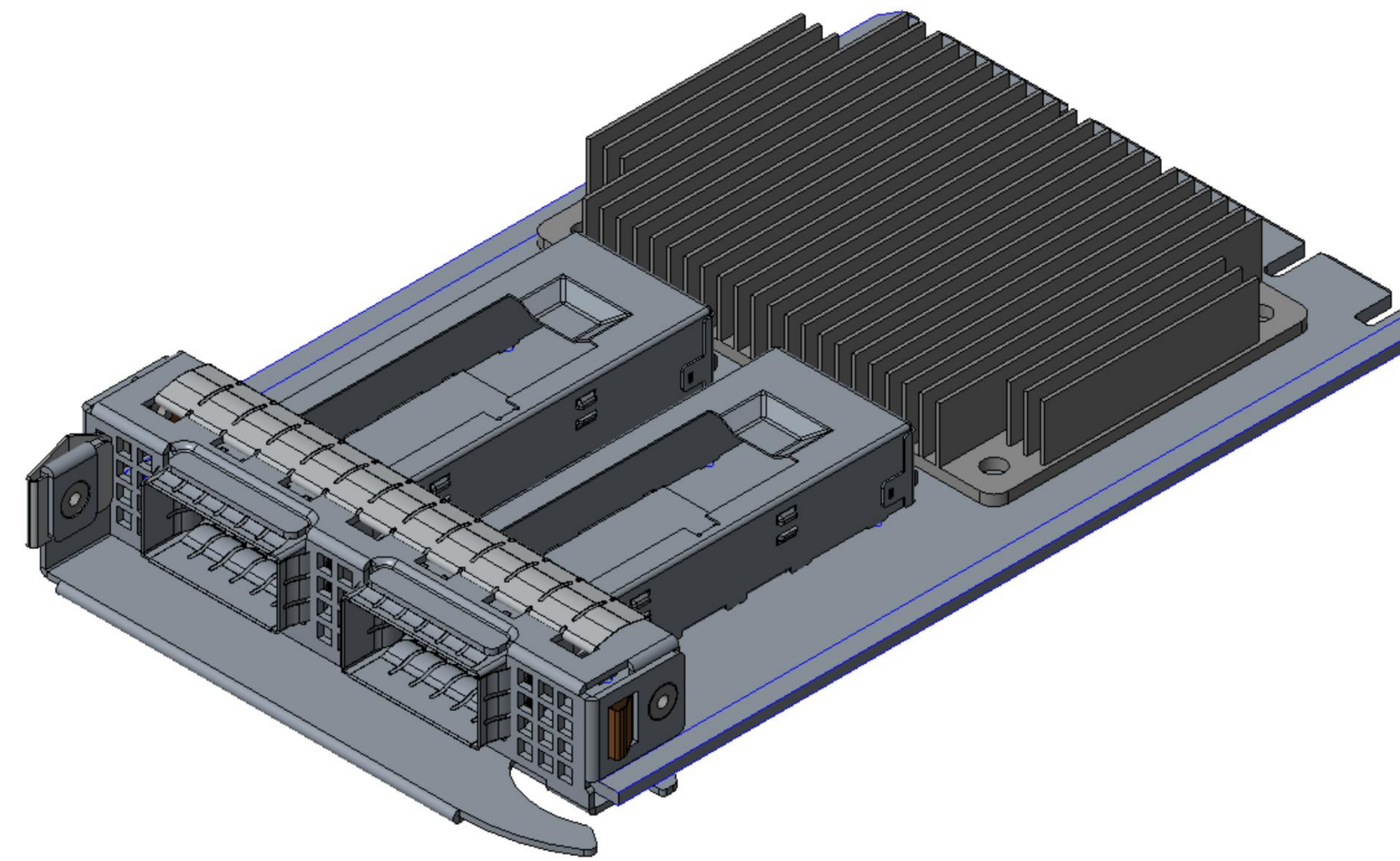
OCP NIC 3.0 Thermal

Thermal Benefits for NIC 3.0

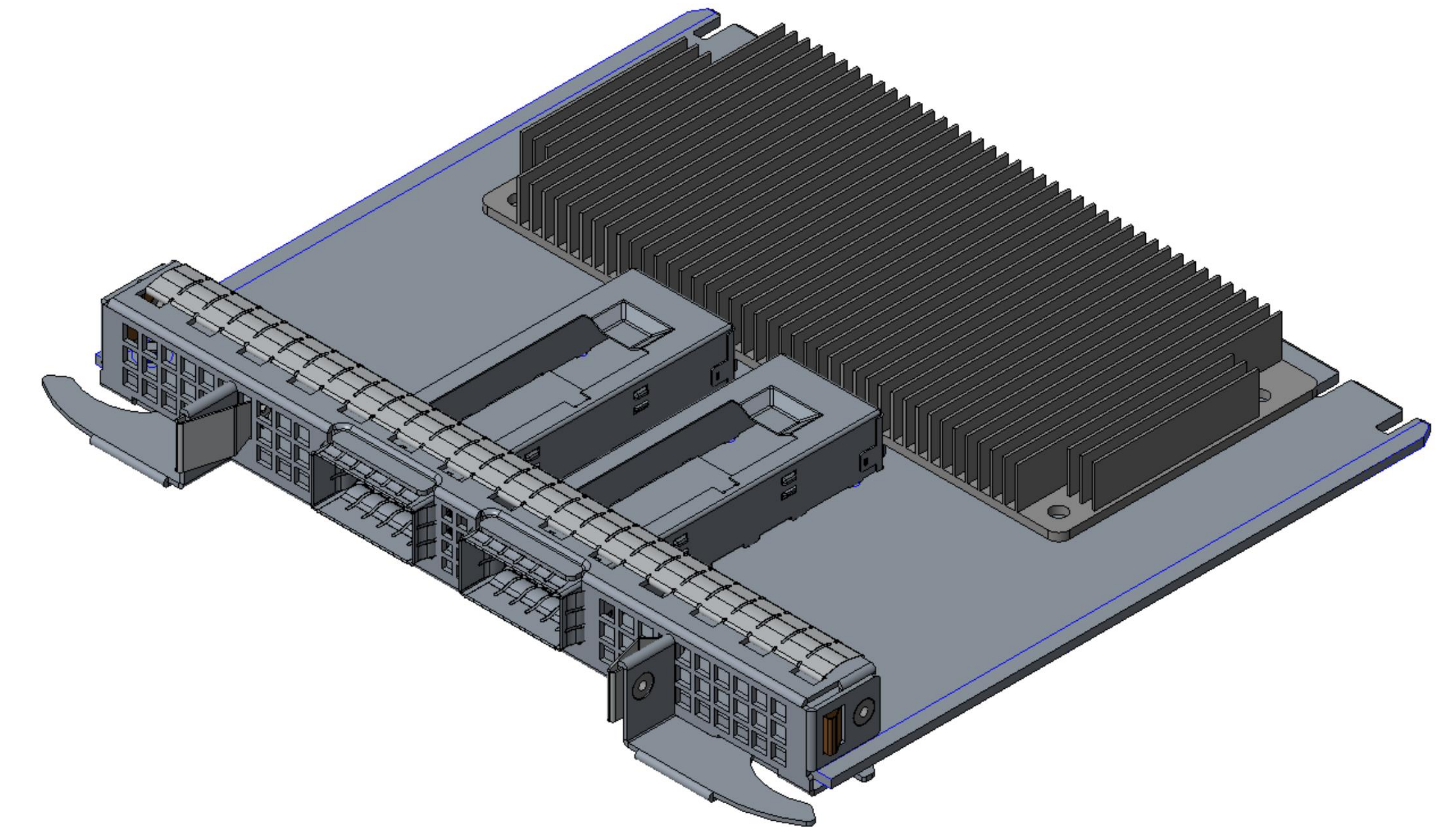
Mezz 2.0



NIC 3.0 SFF



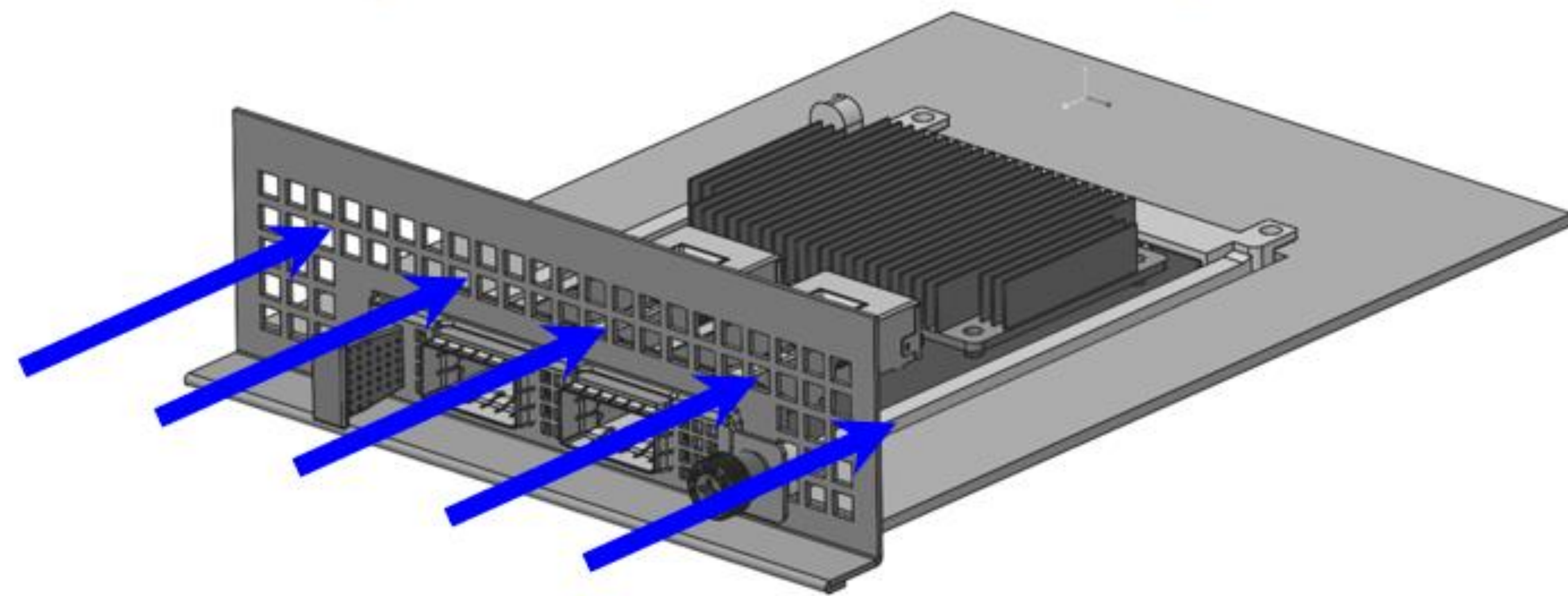
NIC 3.0 LFF

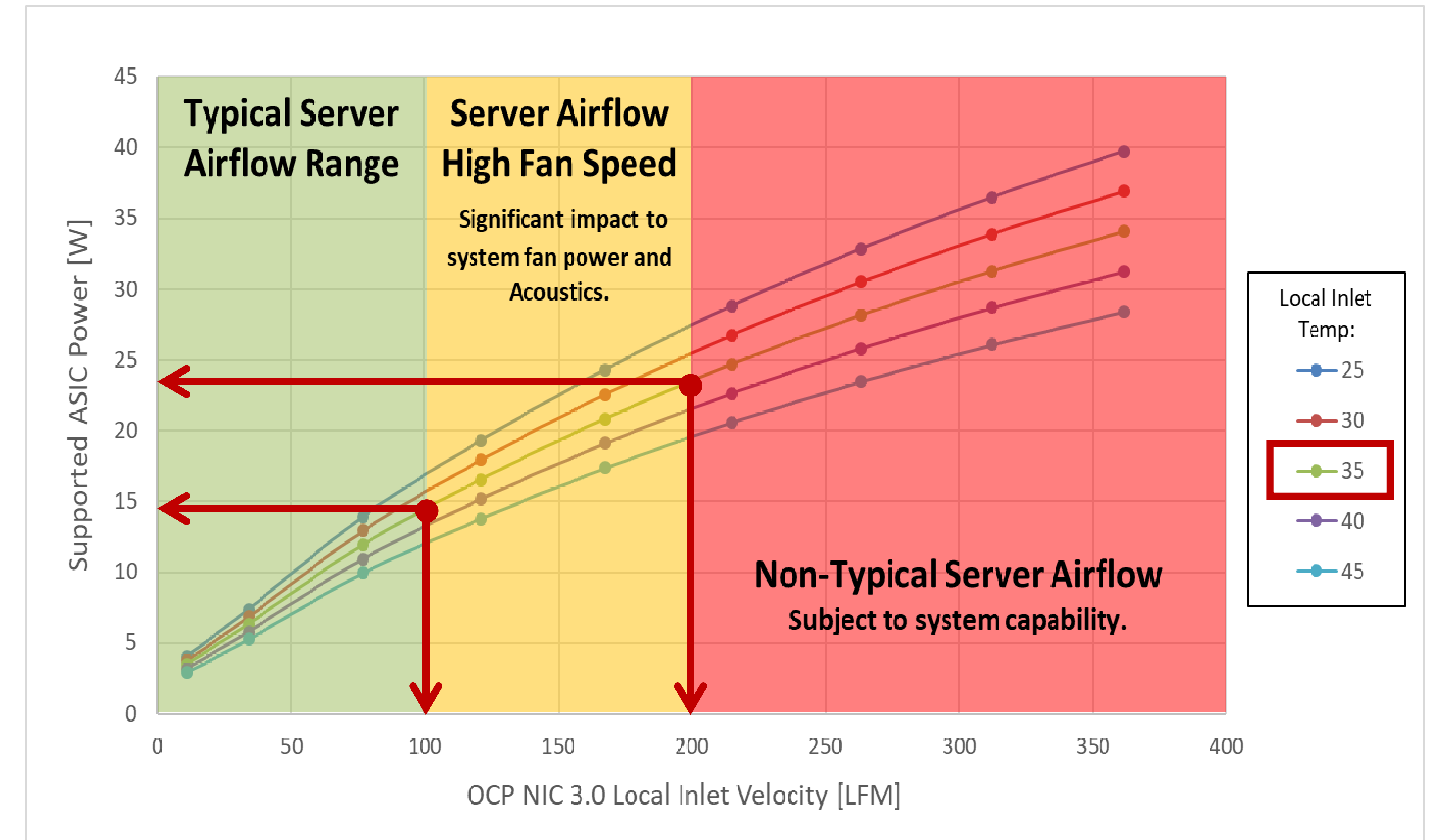


- More space for heatsink with no bergstak connector on the side
- Up-facing heatsink permits flexibility on heatsink height
- LFF allows thermal potential for high-power ASIC cooling

Complete CFD model available at:
http://www.opencompute.org/wiki/Server/Mezz

Thermal Design Guidance – Cold Aisle

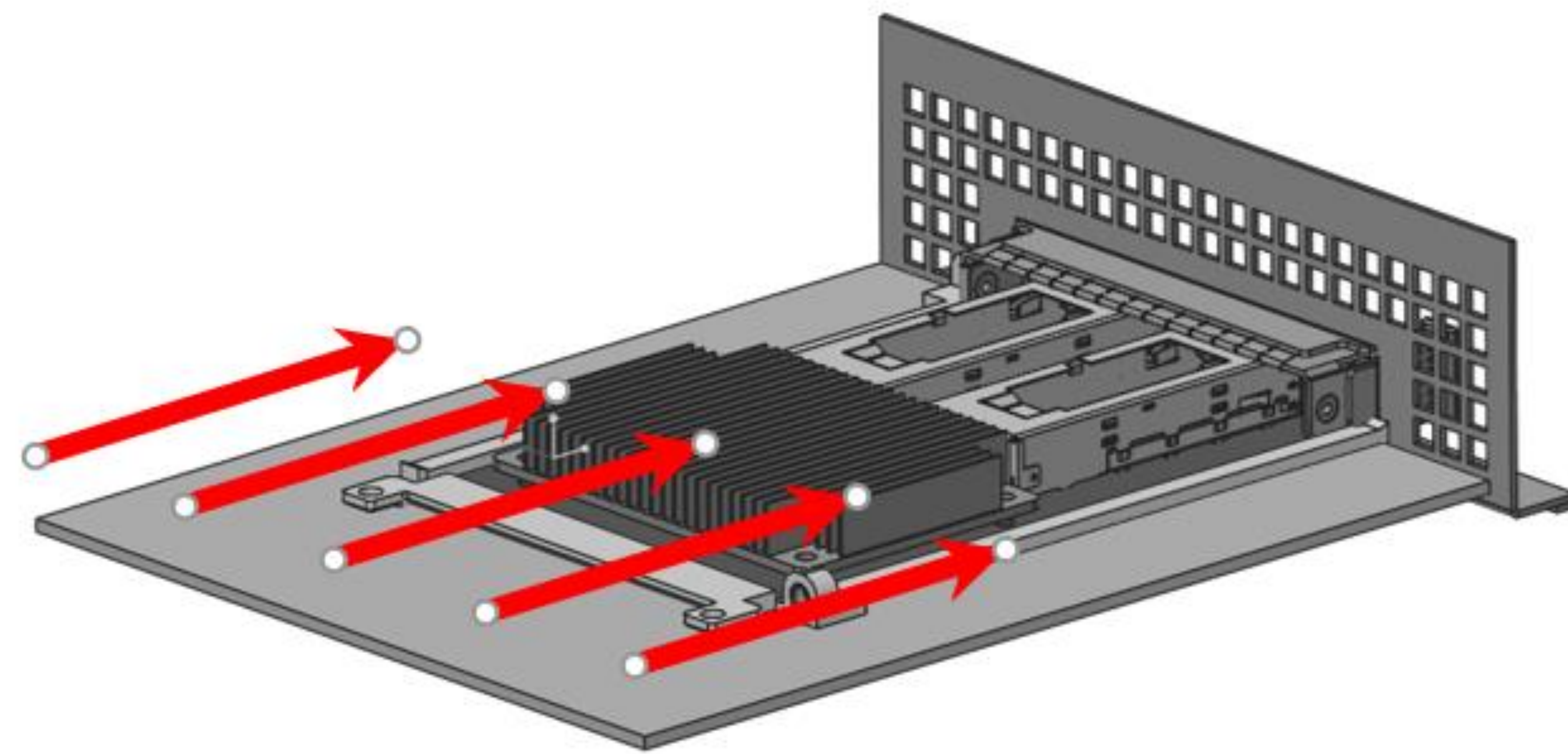


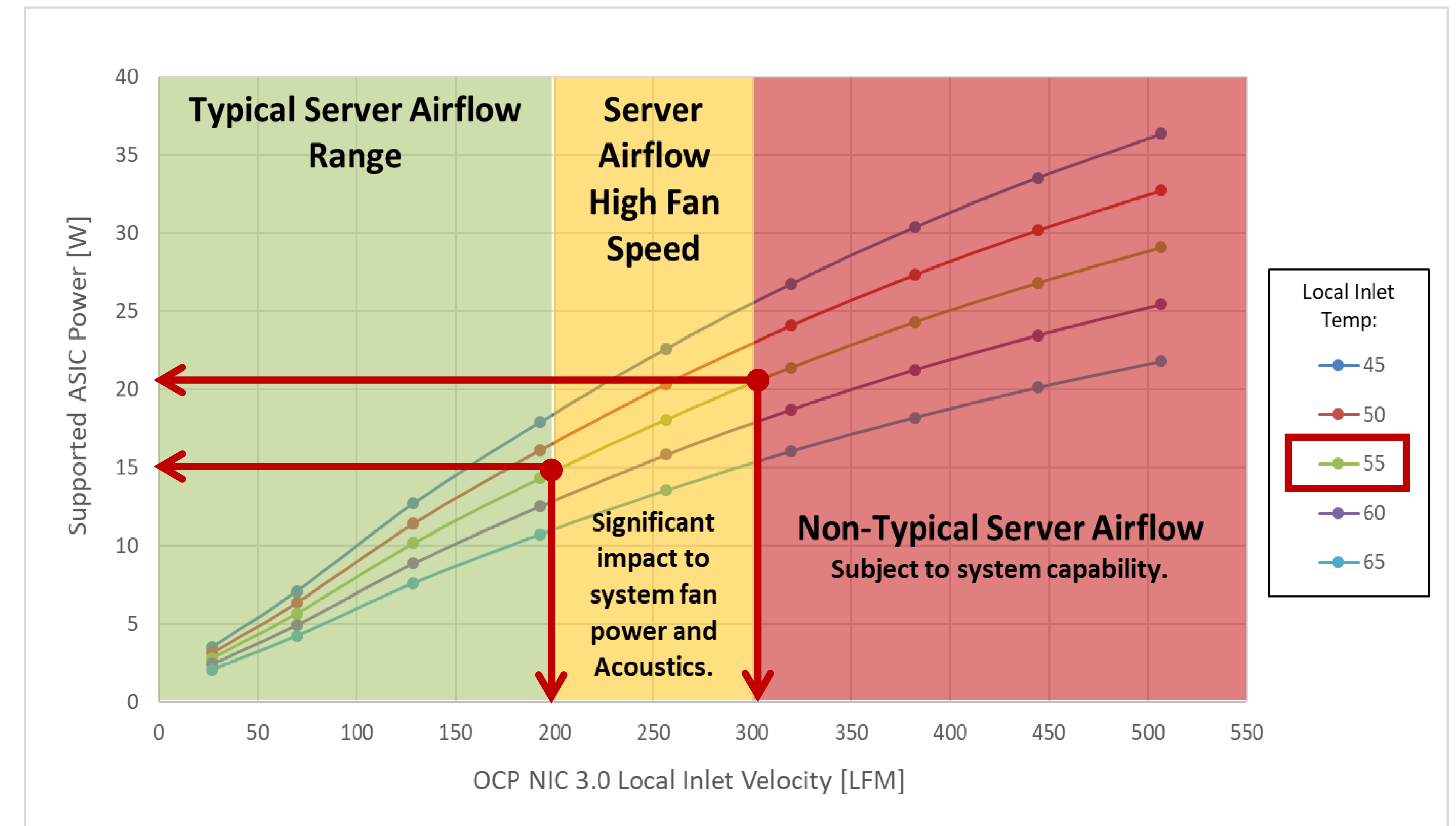


- For typical inlet temperature 35°C, SFF:
  - Support 15W ASICs under 100 LFM (Typical)
  - Support 23W ASICs under 200 LFM (High)

Thermal Design Guidance – Hot Aisle

Complete CFD model available at:
http://www.opencompute.org/wiki/Server/Mezz





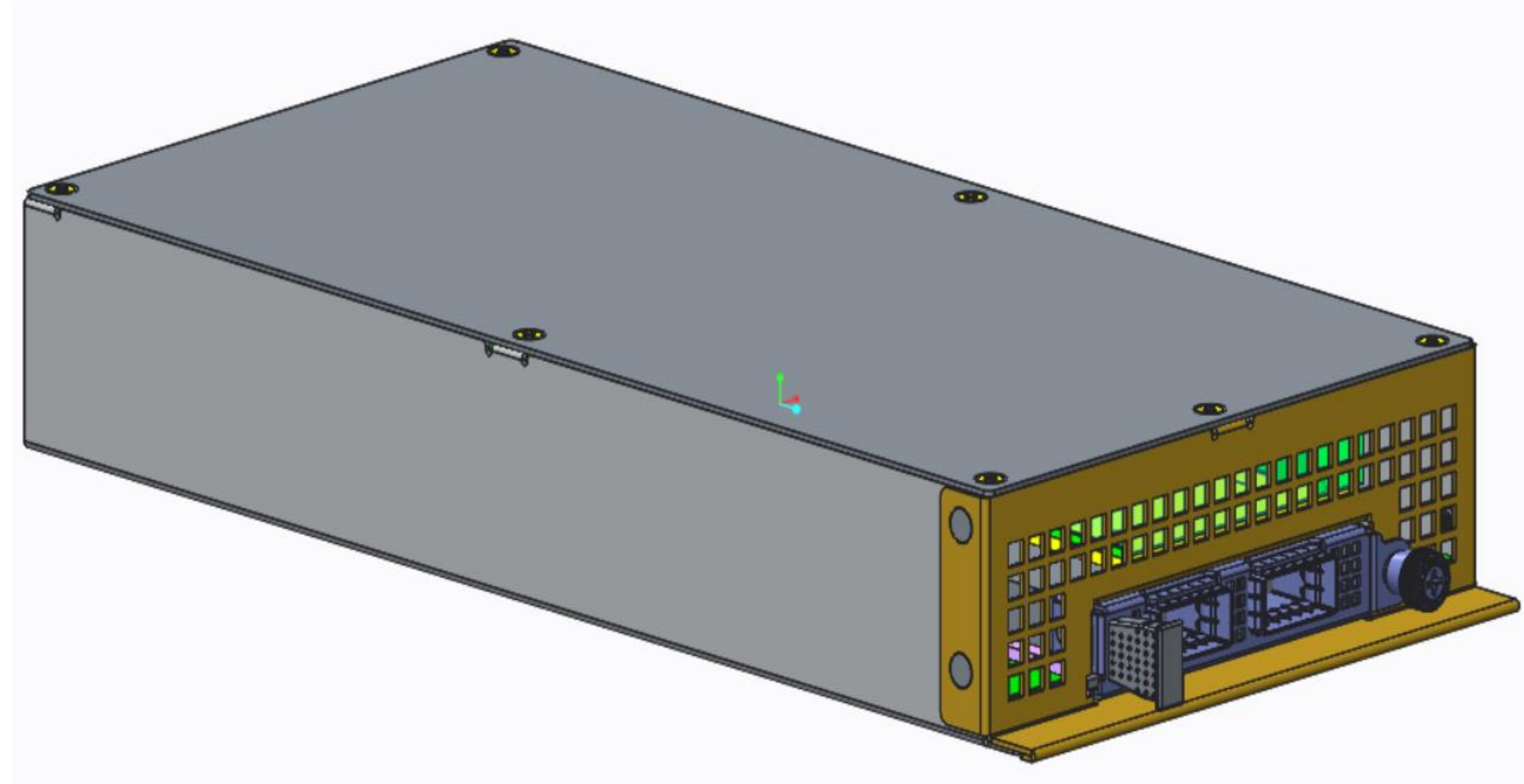
- Hot-aisle cooling is more challenging due to higher supply air temperature
- For typical inlet temperature 55°C, SFF:

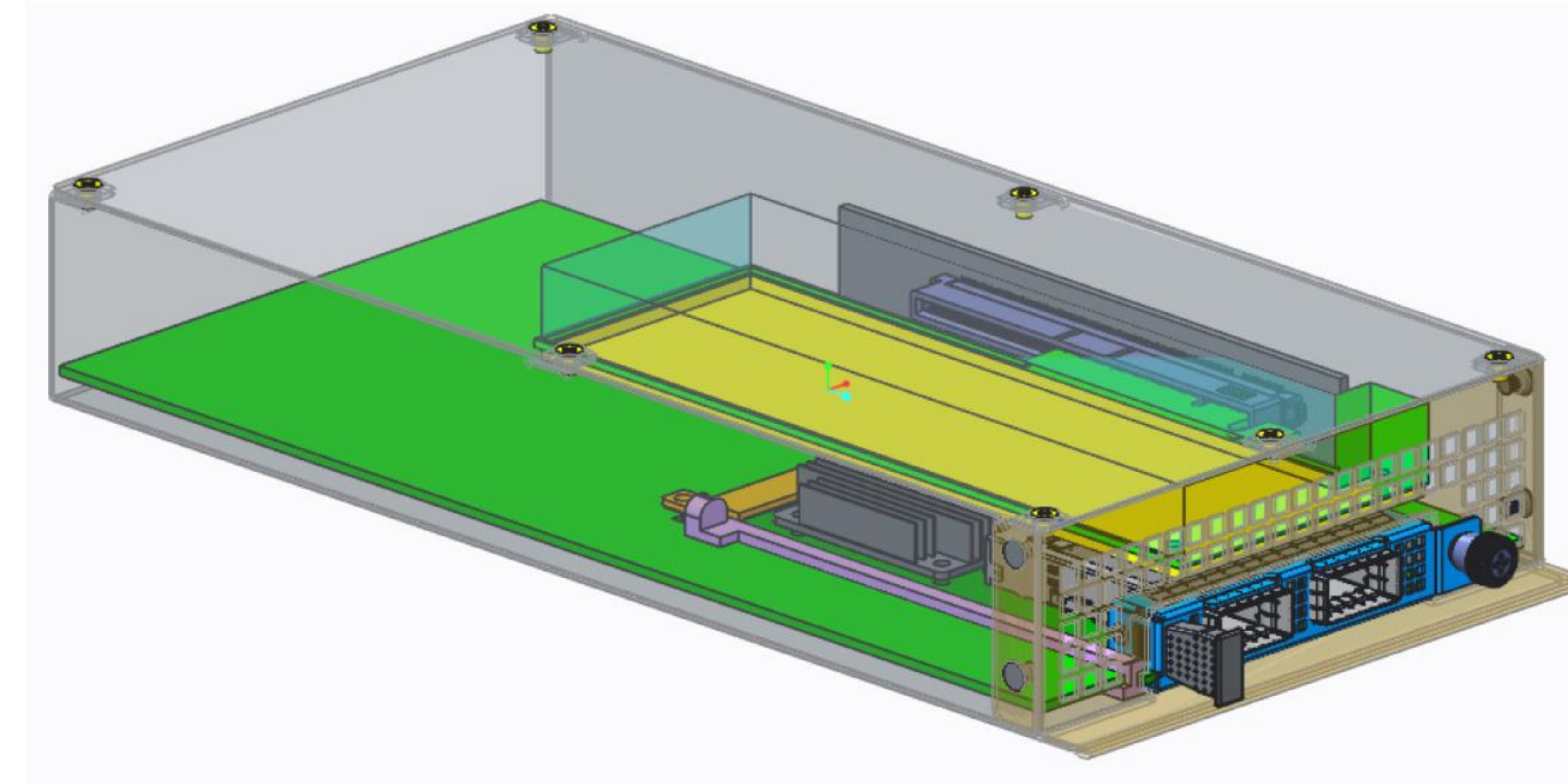
  - Support 15W ASICs under 200 LFM (Typical)

  - Support 20W ASICs under 300 LFM (High)

Preliminary 3D CAD available at:
http://www.opencompute.org/wiki/Server/Mezz

Thermal Test Fixture





- **Purpose:**
  - Provide standardized test data across different NIC and system vendors

- **Features:**
  - Simple and easy adoption by both NIC and system vendors
  - Representative thermal data to define cooling tiers across different use cases
  - Functional test board for power delivery and reporting interface

Implementation and Refinement

0v70 – Initial release
Jan 25th ,2018

0v80 – Hot fixes (Minor mechanical spec updates)
End of Feb’18

0v90 – Add signal integrity guideline and conformance
TBD

1v00 – Add Implementation learning (planned mechanical expansion)
Q4’18

Subgroup Wiki with latest specification : http://www.opencompute.org/wiki/Server/Mezz

Mailing list: http://lists.opencompute.org/mailman/listinfo/opencompute-mezz-card

OCP SUMMIT