OPEN POSSIBILITIES.

OCP NIC 3.0: Thermal Considerations

OCP GLOBAL SUMMIT
NOVEMBER 9-10, 2021

OCP NIC 3.0: Thermal Considerations

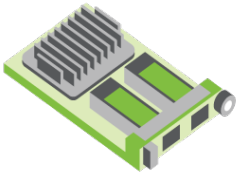
Anna Sokolovskaya, Thermal Engineer, Nvidia
Hasan Ali, Thermal Engineer, Molex
Yin Hang, Thermal Engineer, Meta

OPEN PLATINUM™

OPEN POSSIBILITIES.

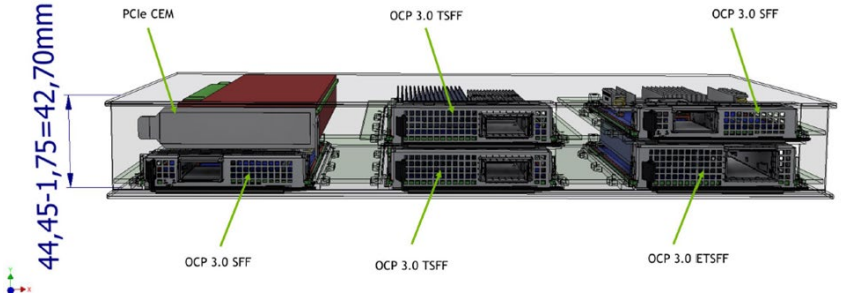
OCP GLOBAL SUMMIT
NOVEMBER 9-10, 2021

Technical Overview



NIC3.0

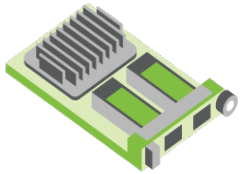
- 400Gb/s NIC → 200G to 400G per port
- Serviceability challenges with DAC cables → Higher power AOCs
- Higher power AOCs → viability of existing infrastructure
- Each °C can have an impact on conclusion
- Alignment on boundary conditions → representative of realistic scenario
- SFF, TSFF (thermal tradeoffs)



OPEN POSSIBILITIES.

OCP GLOBAL SUMMIT
NOVEMBER 9-10, 2021

Focus Areas



NIC3.0

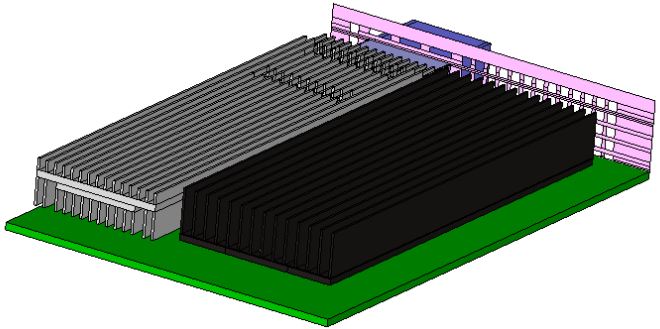
- Form factor: TSFF vs. SFF
- NIC ASIC power limitation: DAC cable only
- Type of module: QSFP-DD Type 1 and Type 2A
- Monitor location: Mean top BS temp, heatsink base temp, nose temp
- Type of testing fixture: with and without testing fixture
- Cold aisle vs. hot aisle

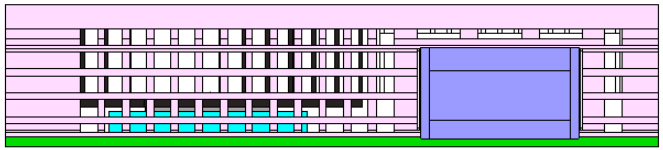
Model Setup

- OCP3 card TSFF (and SFF)

- ASIC: ConnectX-6Dx (Nvidia design)
- ASIC power: 23W
- ASIC HSK: 94mm X 35.5mm X 11mm (8.7mm for SFF)
- Fins thickness: 0.5mm
- Pitch: 2.7mm
- Material Aluminum

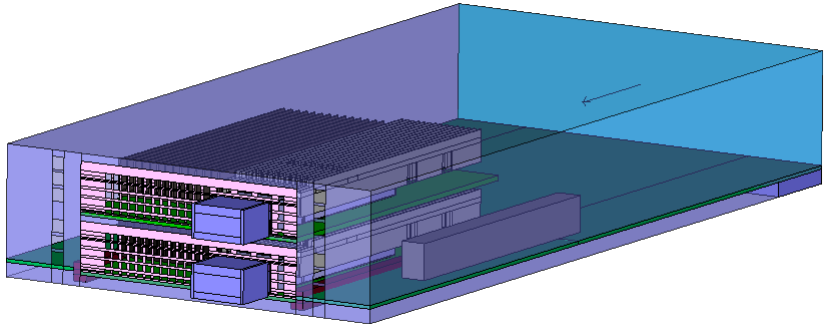
- QSFP-DD - Molex Multilane thermal model
- QSFP-DD power: 10.2W
- QSFP-DD HSK - Molex design





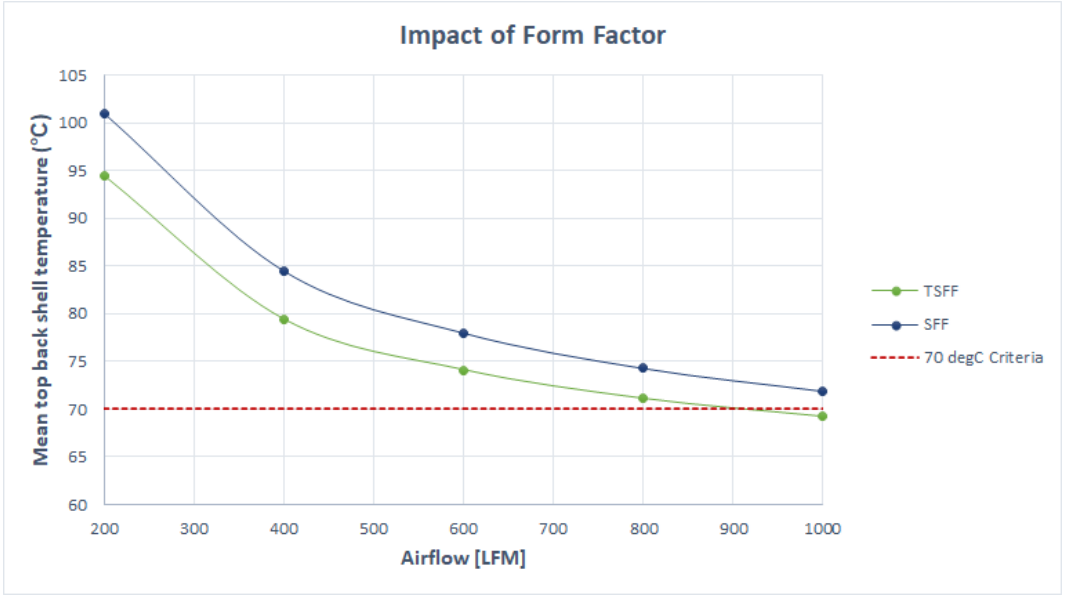
Simulation assumptions

- Hot-aisle
  - Ambient temperature: 55°C
  - Velocity range: 200-1000 LFM
  - Airflow direction: Back to Front
- Cold-aisle
  - Ambient temperature: 35°C
  - Velocity range: 200-600 LFM
  - Airflow direction: Front to Back
- Test fixture: Nvidia OCP3 test fixture
  - 2 identical cards are installed inside the test chamber
  - Only bottom card is under test
  - Top card is used only as a dummy card



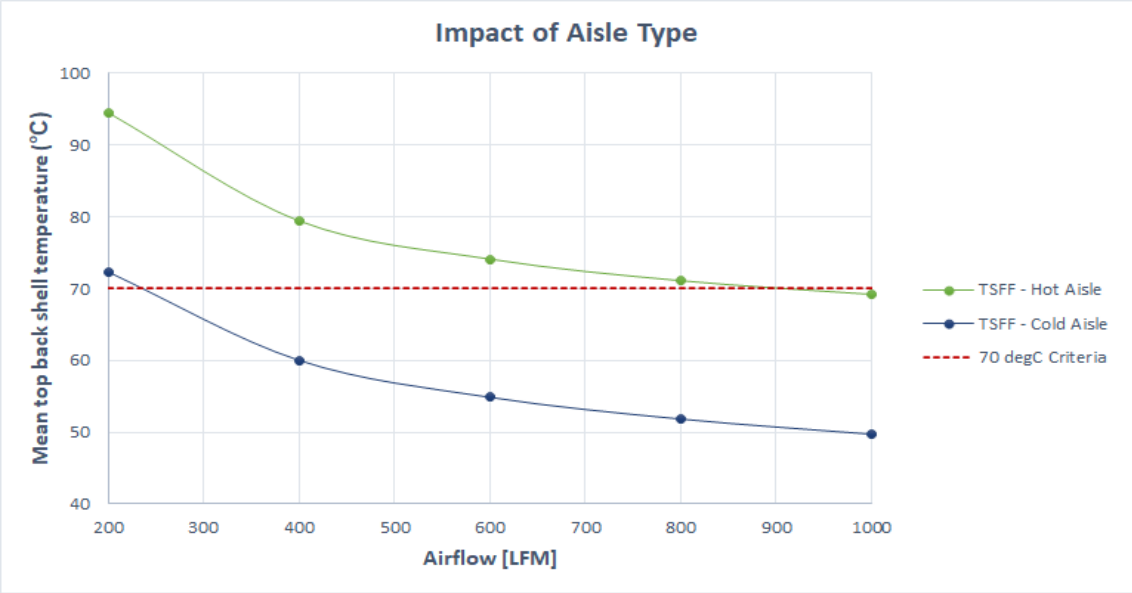
Form Factor

- TSFF has significantly better thermal performance owing to additional space for increased fin height along with added ventilation to make the QSFP-DD heatsink more effective.
- At lower airflow, TSFF can improve thermal performance by as much as 6°C.

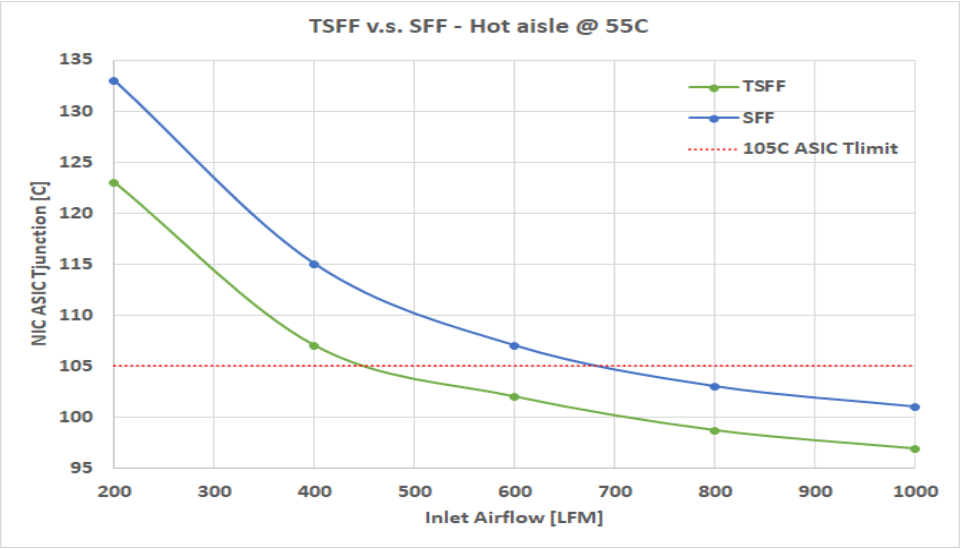


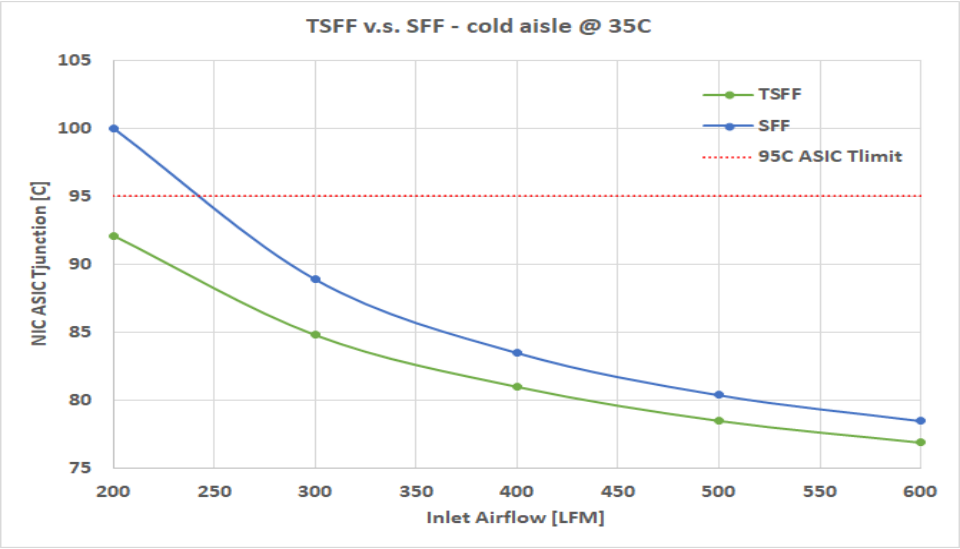
TSFF - Aisle Comparison

- With 20°C lower ambient and front to back airflow, mean top back shell temperatures are significantly lower in the cold aisle.



Power Limitations on NIC ASIC



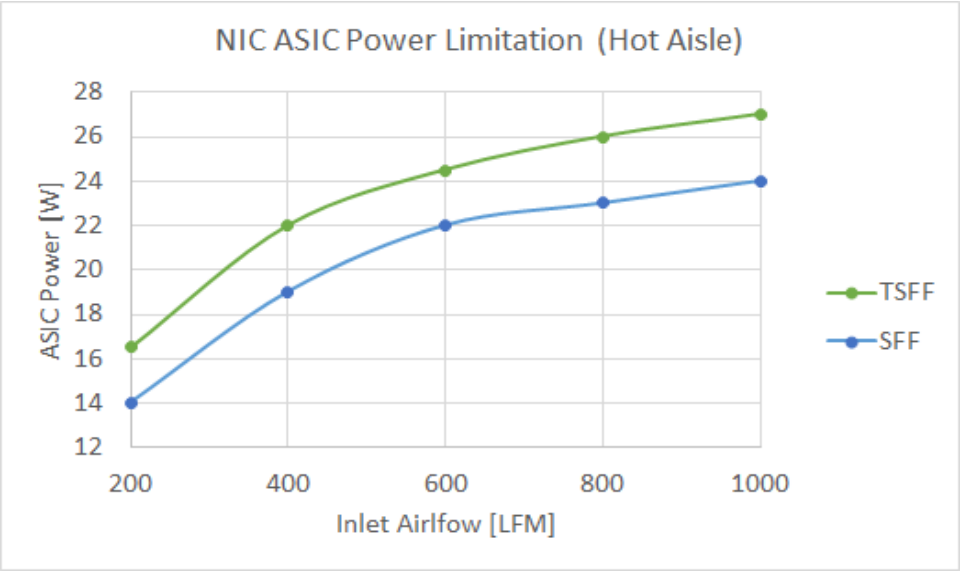


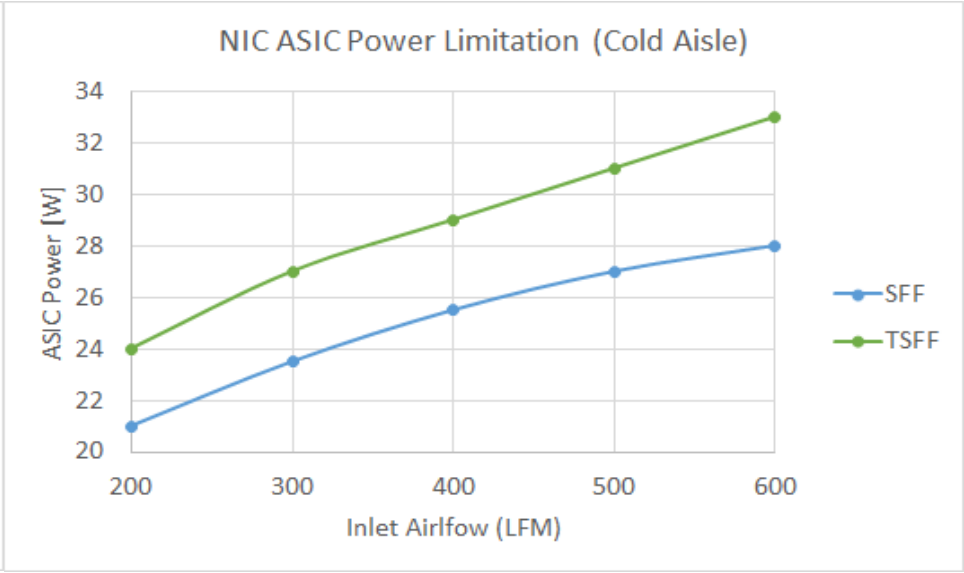
- 23W NIC ASIC power and DAC QSFP-DD is assumed.
- TSFF improves the thermal performance by 4-10°C for the hot aisle side and 2-8°C for the cold aisle side. The thermal performance gains get more significant when airflow is low.

OPEN POSSIBILITIES.

OCP GLOBAL SUMMIT
NOVEMBER 9-10, 2021

Power Limitations on NIC ASIC





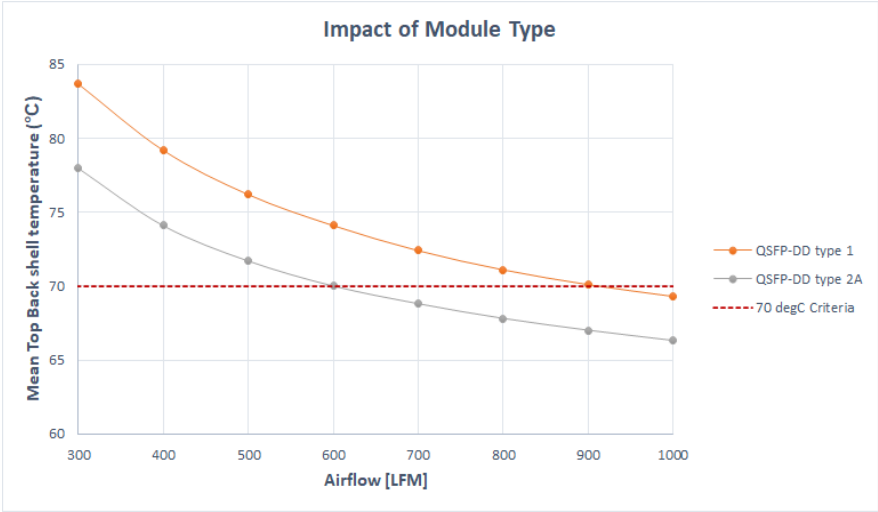
- Hot aisle assumes 105°C as the Tj limitation and cold aisle assumes 95°C as the Tj limitation per application experience.

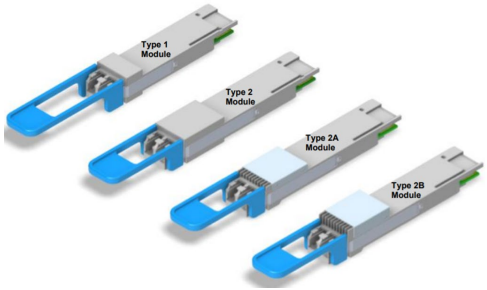
OPEN POSSIBILITIES.

OCP GLOBAL SUMMIT
NOVEMBER 9-10, 2021

Type of Module

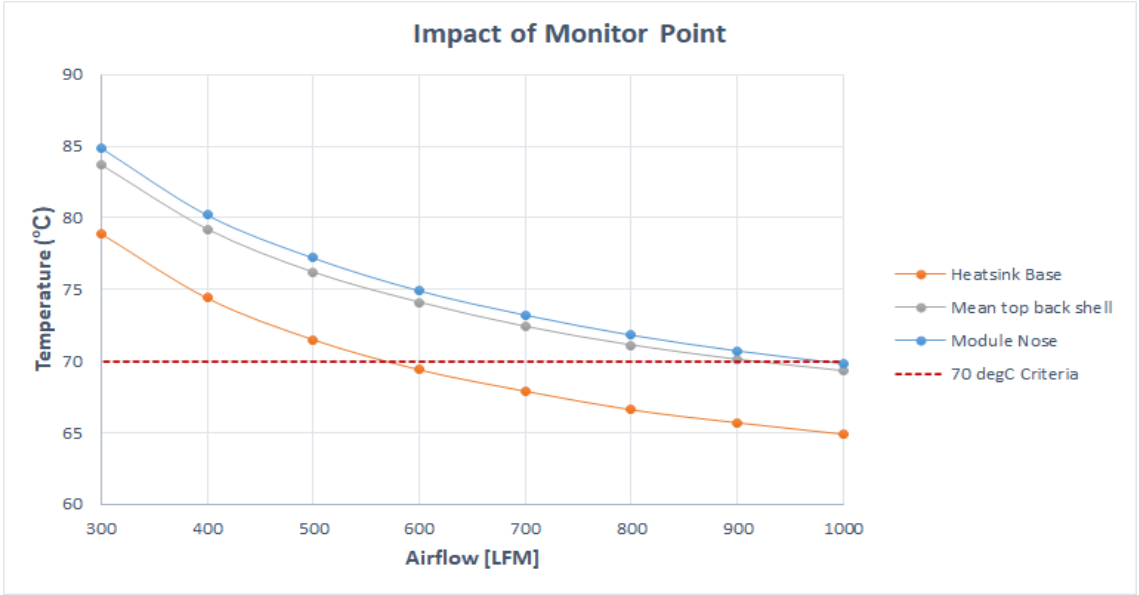
- Type 2A QSFP-DD has better performance than type 1 due to its external integrated heat sink in the rear part of the cable.
- It's important to understand what QSFP-DD module type is being tested while we discuss thermal performance of the module or heat sink thermal design.

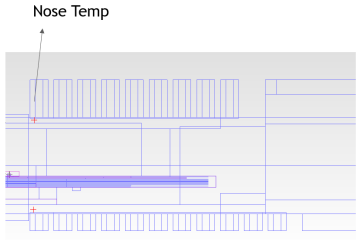




Monitor Location

- When the thermal performance of the cable is discussed, it's important to understand what is the monitor point and the criteria for pass/fail.
- Different monitor points have different temperatures, and the differences can be very significant.
- Normally we see the nose temp should be lowest, for this presentation, we use a power profile which has power dissipated at the nose area.



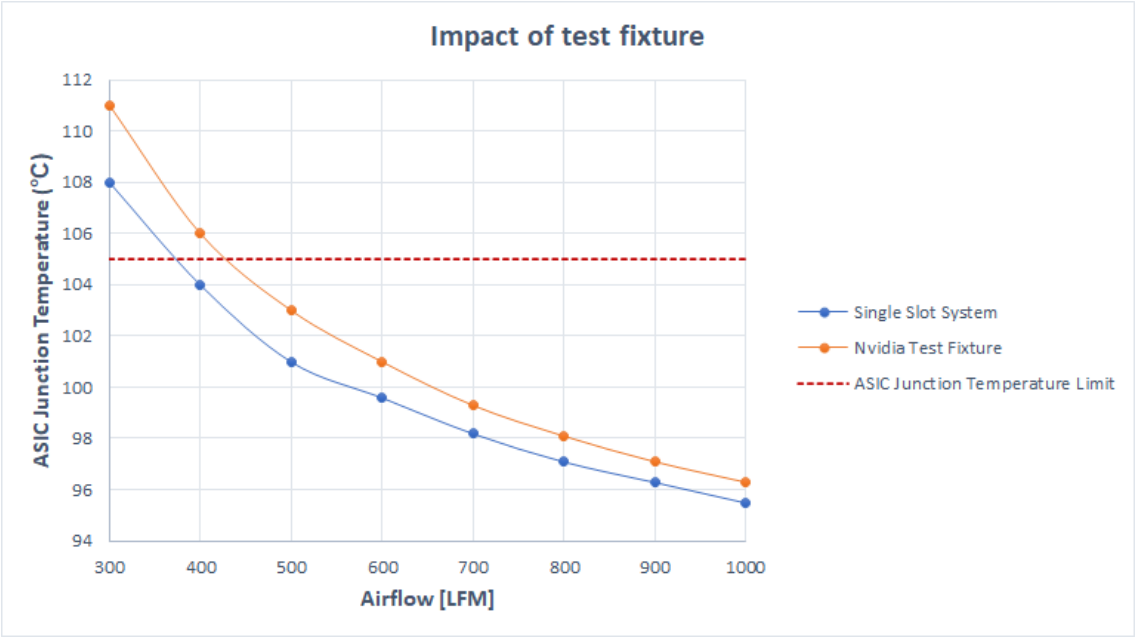


OPEN POSSIBILITIES.

OCP GLOBAL SUMMIT
NOVEMBER 9-10, 2021

Test Fixture

- Single slot system has better thermal performance (by 1-3°C, depending on airflow) compared to NIC simulated in the test fixture.
- Test fixture more accurately represents air bypass in real applications.



Call to Action

- Significant impact of variables and boundary conditions
- Consensus needed regarding to the monitoring temperature location
- Active participation from module, IO, NIC, system and data center architects
- Expand study to OSFP-RHS, QSFP-DD AEC
- NIC project Wiki with latest specification: http://www.opencompute.org/wiki/Server/Mezz
- Mailing list: http://lists.opencompute.org/mailman/listinfo/opencompute-mezz-card

Thank you!

OCP GLOBAL SUMMIT

NOVEMBER 9-10, 2021