Modular Heterogeneous Compute Platforms using Intel CPU and NVIDIA GPGPU modules

Concurrent Technologies – Our Mission

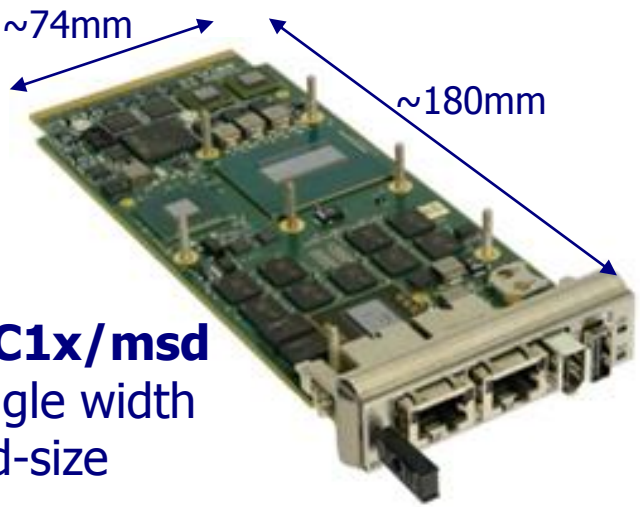
- To design and manufacture high performance, modular and scalable computing solutions based on open standards
- To provide application enabling middleware that shields the user from underlying complexity

- Main markets today include:
  - Defense and Aerospace
  - Industrial Automation
  - Telecommunications
  - Transportation
  - Space and Energy Research

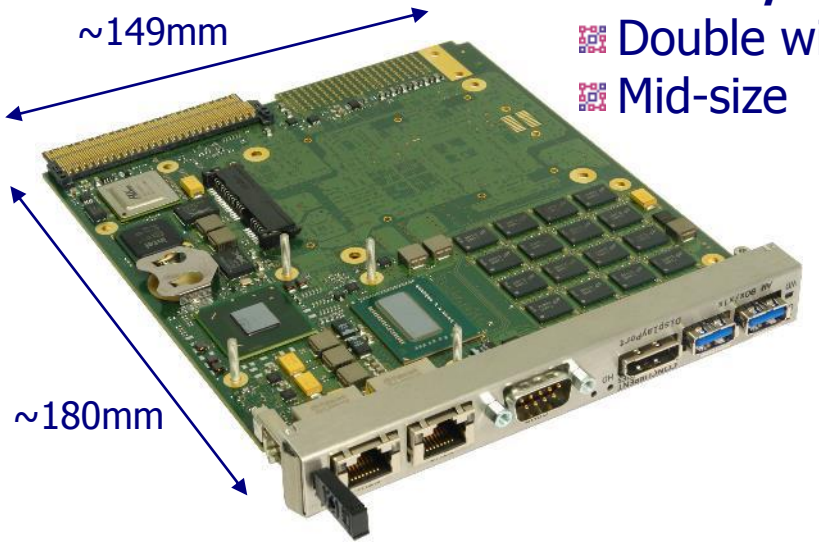
AdvancedMC (AMC)Form Factors

- **AMCs are available in two widths:**
  - Single, approximately 74mm
  - Double, approximately 149mm
- **There are three AMC module sizes:**
  - Compact, approximately 13mm high
  - Mid-size, approximately 18mm high
  - Full-size, approximately 28mm high

} Processor modules are typically Mid and Full-size



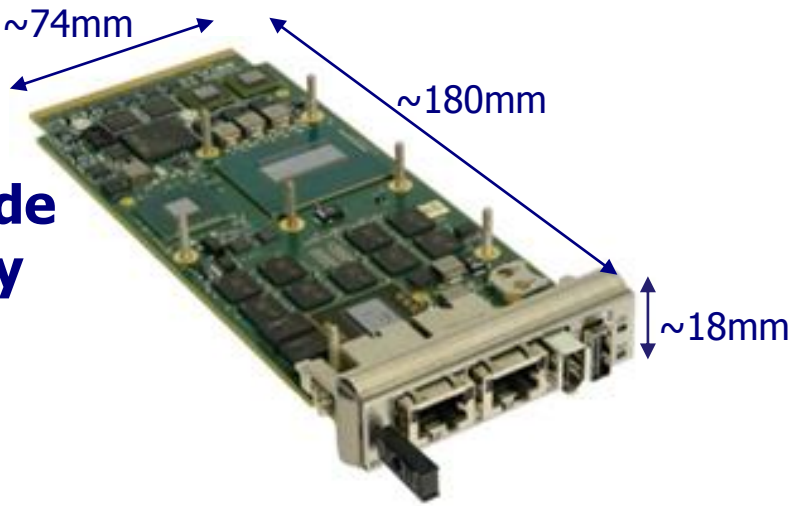
AM C1x/msd
- Single width
- Mid-size



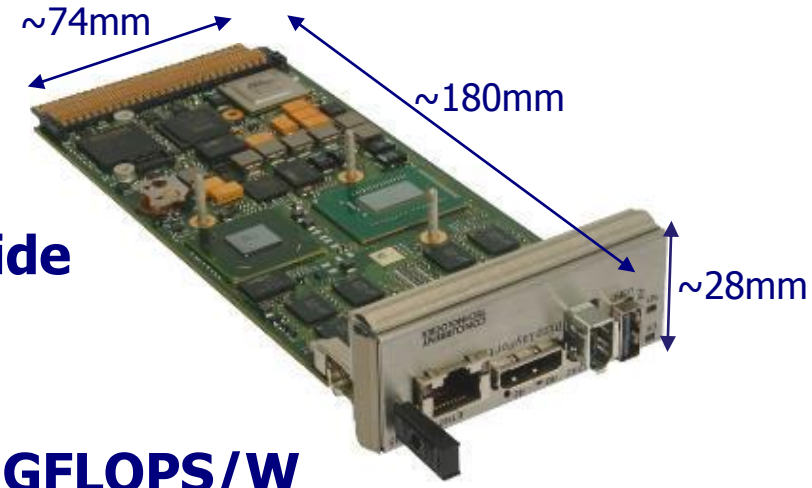
AM 902/411
- Double width
- Mid-size

Basic unit of computing

- **Mid Size:**
  - 74 x 180 x 18 = 239760mm^3
  - ~ 240cm^3
- **Within this envelope we can provide 650 GFLOPS processing capability**



- **Full Size:**
  - 74 x 180 x 28 = 372960mm^3
  - ~ 373cm^3
- **Within this envelope we can provide 1.3 TFLOPS processing capability**



- **Theoretically this equates to ~32 GFLOPS/W**
- **But how do they scale to HPC levels…..**

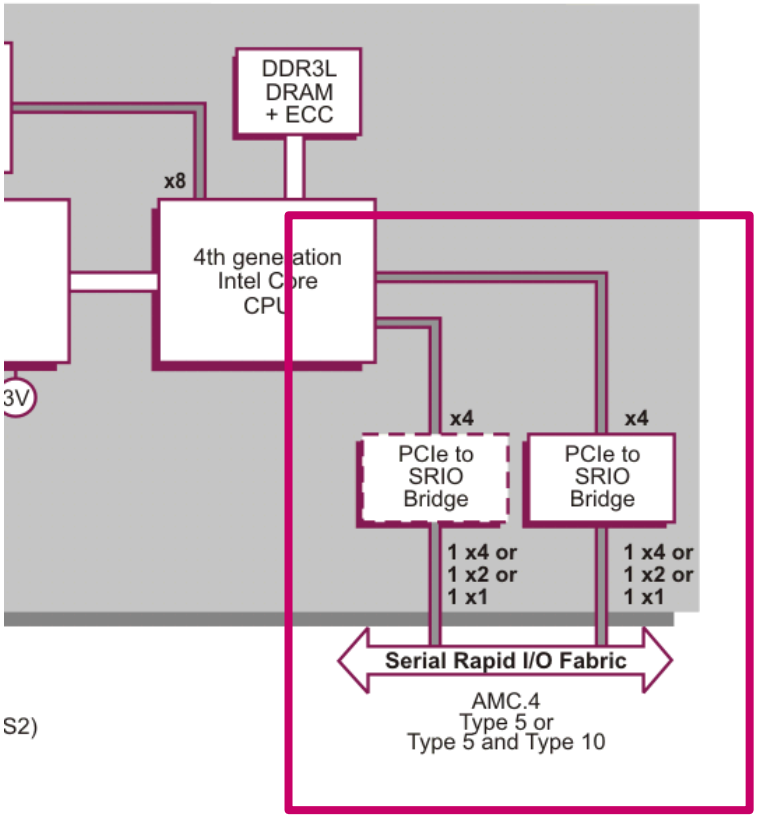
How do we scale these modules?

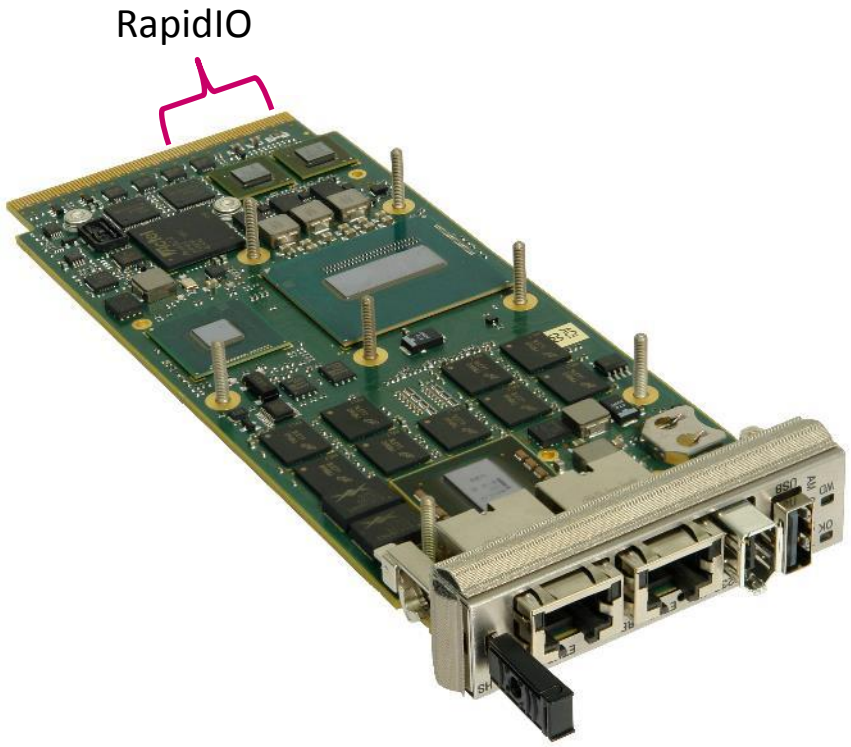
RapidIO Interconnect 'Check In'

Interconnect Requirements	RapidIO	Infiniband	Ethernet	PCIe	The Meaning of ●
Low Latency	●	✗	✗	●	switch silicon: ~100 nSec memory to memory : < 1 uSec
Scalability	●	●	●	✗	support any topology, 1000's of nodes, true peer-to-peer
Integrated HW Termination	●	✗	✗	●	integrated into SoCs and guaranteed, in order delivery without software overhead
Power Efficient	●	✗	✗	●	3 layers terminated in hardware, Integrated into SoC's
Fault Tolerant	●	●	●	✗	supports hot swap and fault tolerance
Deterministic	●	●	✗	●	guaranteed, in order delivery with deterministic flow control
Top Line Bandwidth	●	●	●	✗	supports > 8 Gbps/lane

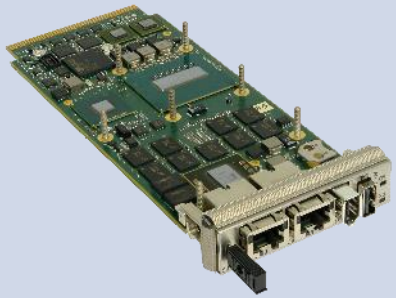
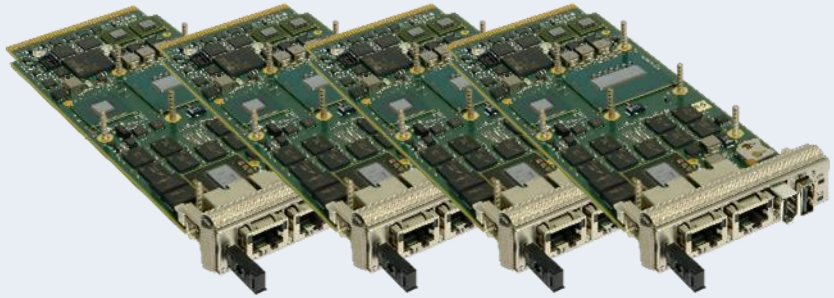
Bridge from native PCIe to RapidIO

RapidIO Bridge on an Intel CPU module



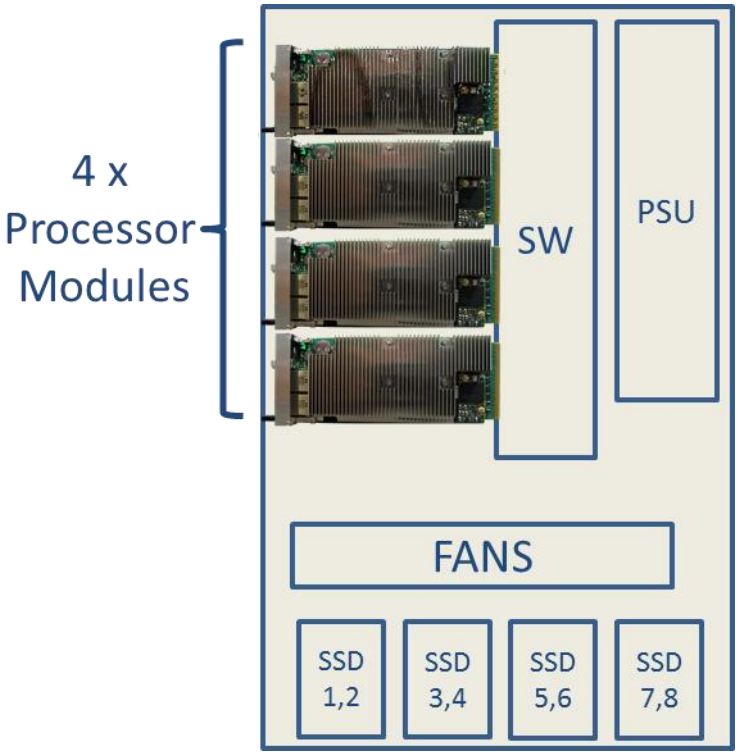


Mix and Match flexibility

Intel CPU Module	NVIDIA GPGPU Module
	
	
	

19” Rack Mount Enclosure

- **Data Centre Compute and Networking (DCCN) platform has been released**
  - Proof of concept enclosure targeted at Server, Data Centre & Supercomputing applications
  - Showcases strength of the RapidIO ecosystem with multi-vendor collaboration

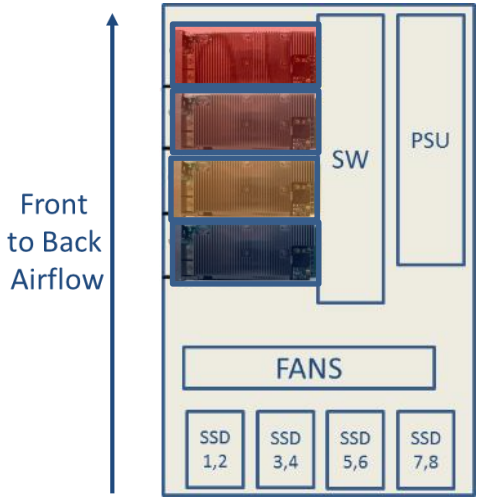


Rear Expansion Connectivity

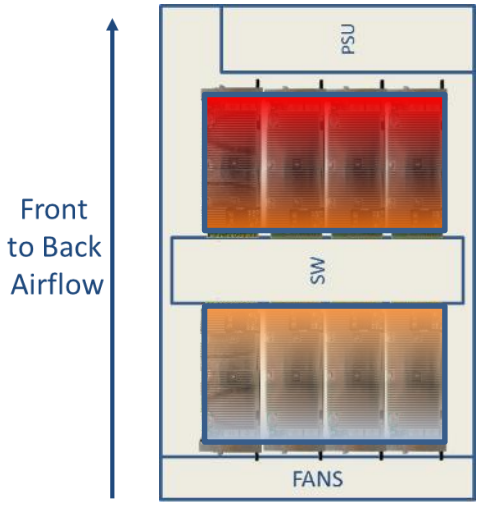
- **Current solution is easy to stack:**
  - 4 x 20 Gbps switched RapidIO
  - 4 x 10 Gigabit Ethernet direct interconnected to AdvancedMC modules
  - 2 x Gigabit switched Ethernet

Future Improvements?

- Current system is not best optimised for cooling
- Air gets successively hotter as it passes over each processor module



- By re-architecting the box it may be possible to double the number of processor nodes



Software Applications

- **Ethernet is the most common interconnect used between servers**
- **Software packages from the HPC space tend to use TCP/IP socket APIs running on a Linux OS:**
  - A good example of this is Hadoop which is an open framework for distributed processing of large data sets across clusters of computers
- **The challenge has been how to utilize this ecosystem of applications in an embedded environment where PCI Express or RapidIO fabric interconnects might be used**

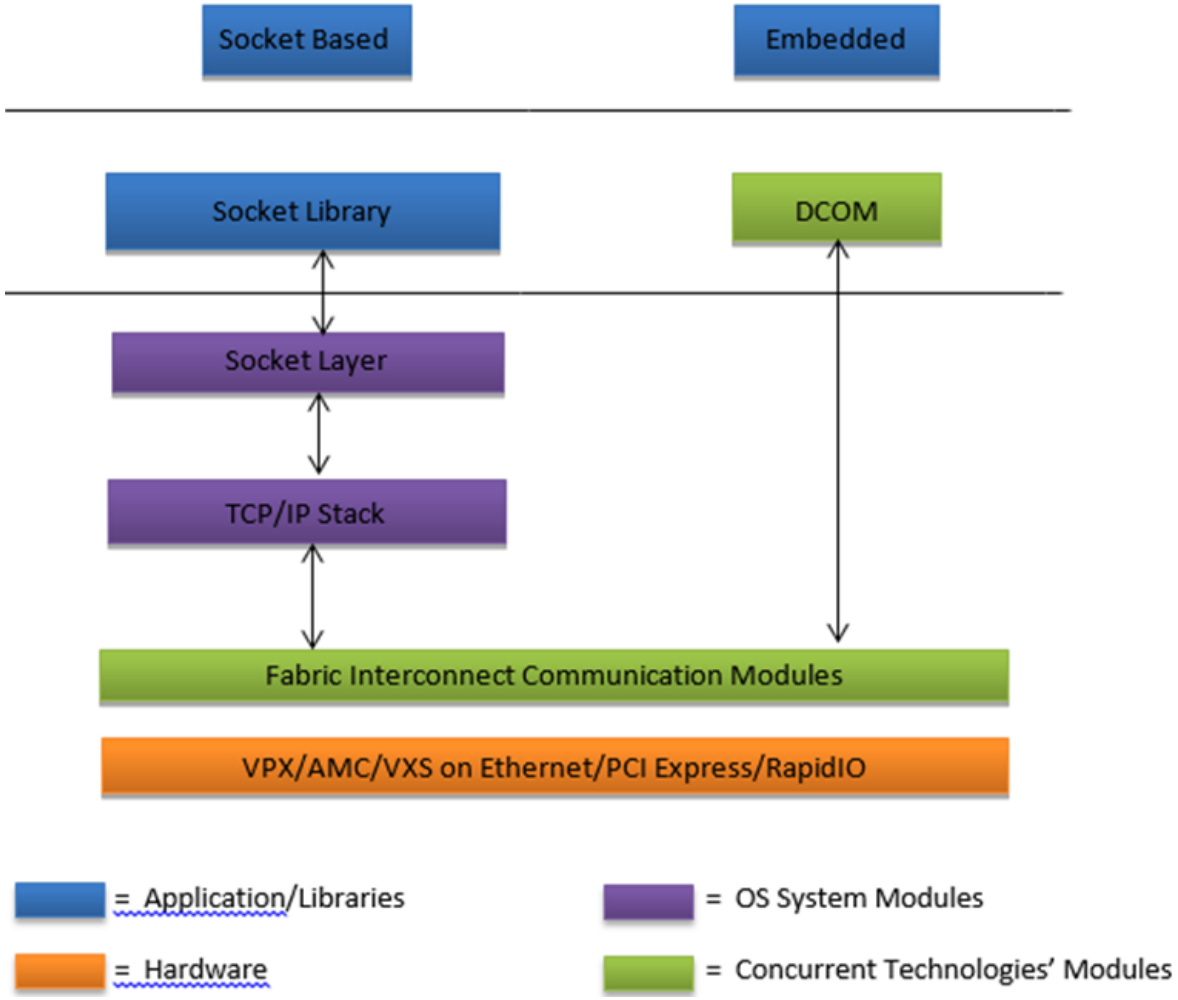
One solution: FIN-S

- **Emulates an Ethernet device over PCI Express or RapidIO**
- **From an application perspective, the interconnect is seen as an Ethernet network running over TCP/IP**
- **FIN-S shields the application from the underlying fabric and allows some useful side benefits:**
  - Improved throughput with PCI Express and RapidIO (slide 14)
  - CPU utilization reduction (slide (15)
  - Best latency with RapidIO (slide 16)

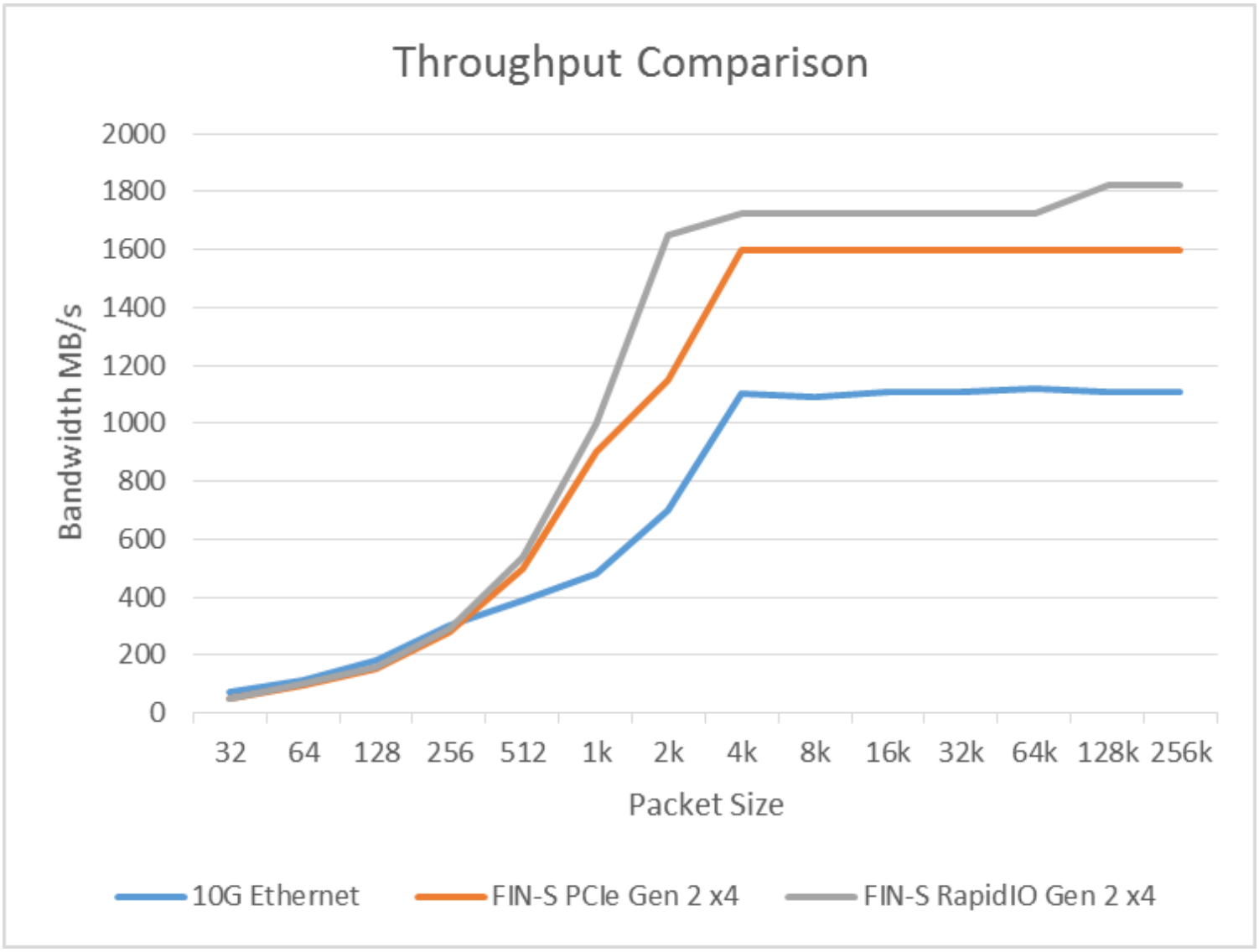
[Comparison was done using:

Processor board with a 10 Gigabit Ethernet adapter connected via a x8 Gen2 PCI Express link
Processor board running FIN-S on a PCI Express Gen2 x4 fabric across the backplane
Processor board running FIN-S on a RapidIO Gen2 (5 Gbps) x4 fabric across the backplane]

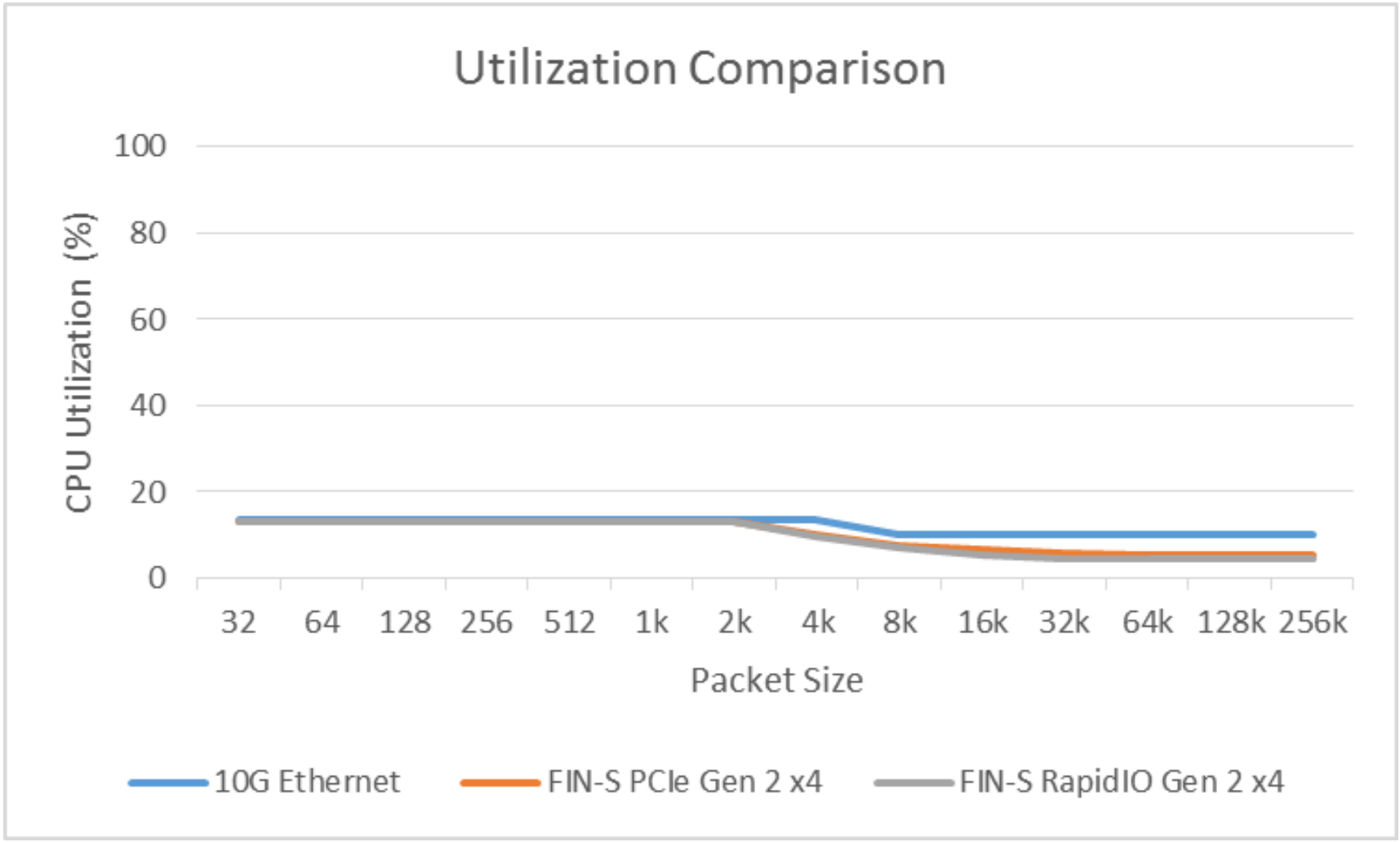
FIN-S Diagram



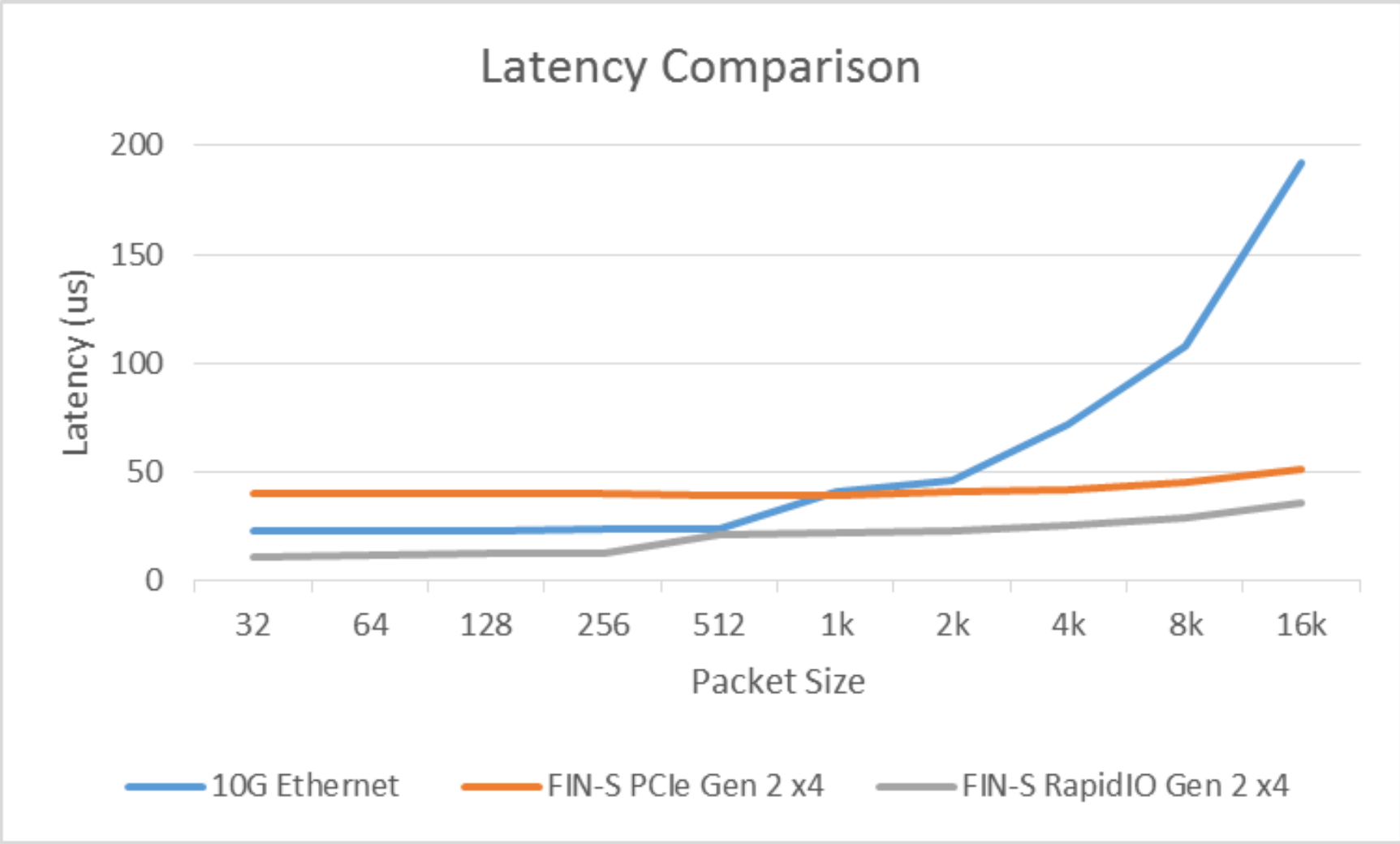
Throughput vs Packet Size



CPU Utilisation vs Packet Size



Latency


Latency Comparison
200
150
100
50
0
Latency (us)
32
64
128
256
512
1k
2k
4k
8k
16k
Packet Size
10G Ethernet
FIN-S PCIe Gen 2 x4
FIN-S RapidIO Gen 2 x4

AdvancedMC based HPC Advantages

- **Key advantage are:**
  - Granularity – small modules are easy to stack
  - Low Latency – RapidIO fabric enables lower system latency for better parallel compute performance
  - Based on Open Standards – strong eco-system and well supported
  - Easy to use – FIN-S layer enables any socket based application to work transparently over RapidIO fabric
  - Scalable up to cabinet level and beyond using top of rack RapidIO switches

Thanks