

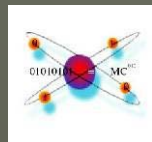
TCP Over PCI Express communication for AdvancedTCA and MicroTCA embedded systems

A software stack IP for low latency-high performance, multi-node,
system to system communication over PCIeexpress Gen2 cable

NucleoDyne Systems, Inc.

19925 Stevens Creek Blvd
Cupertino, California

Contact: ram@nucleodyne.com



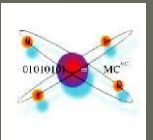
Overview

Traditional Use model

- Ethernet devices connect two or more computer systems.
- PCI Express bus links motherboard mounted peripherals or add-in devices.

New Use Model

- PCI Express switch connects two or more computer systems at system bus level.



Inter process communication schemes

Direct Communication

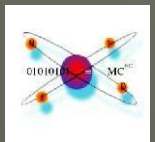
- One process read/writes information to remote process on a different system
 - ❑ No socket API is necessary
 - ❑ Low latency data access

TCP/IP based Communication

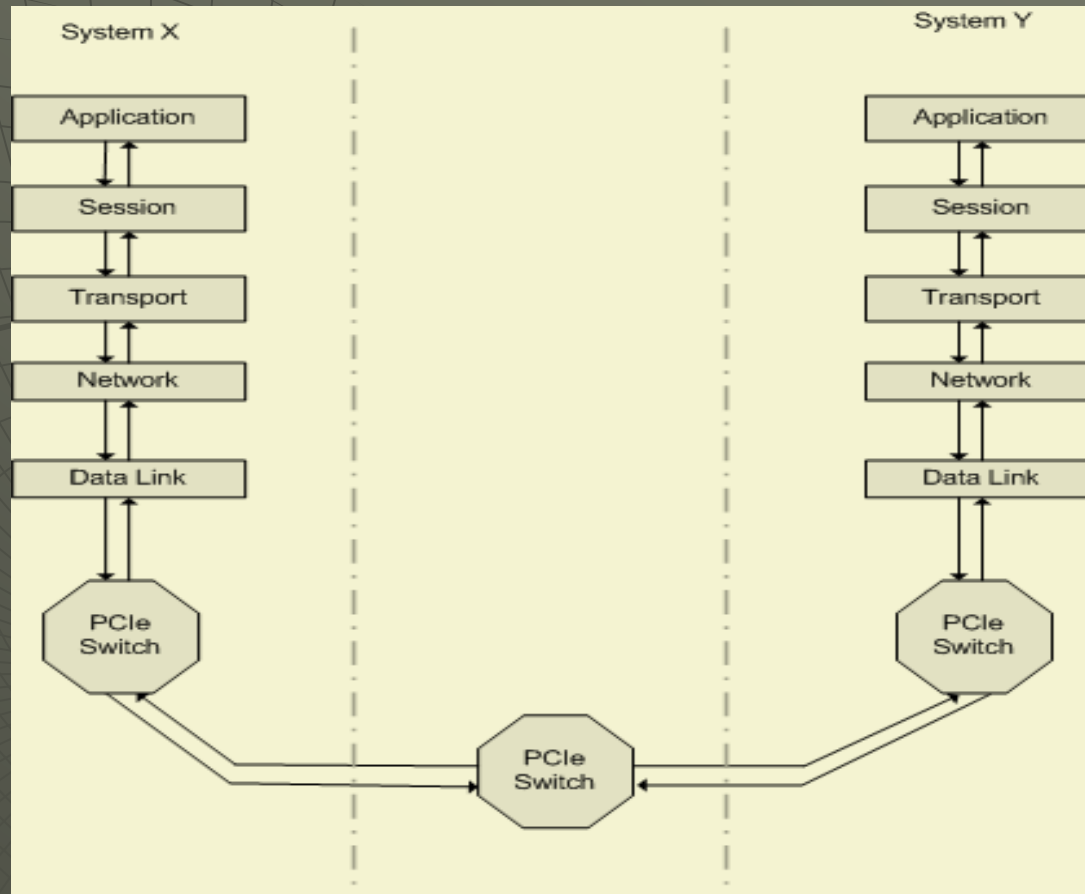
- ❑ Ethernet frames are sent over PCIe Bus

Observations

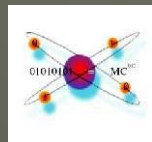
- ❑ Low Latency and High Throughput
- ❑ Lower Power consumption and Low cost



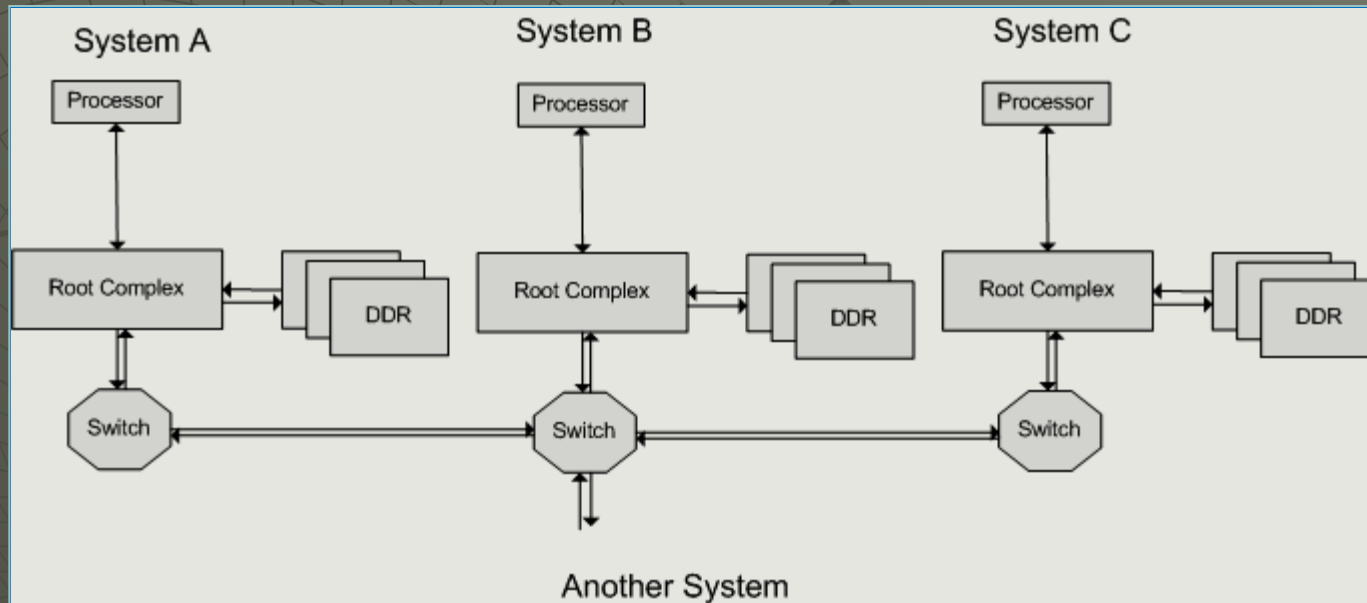
System to System Communication Ethernet Over PCI Express



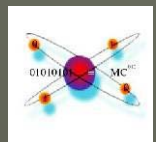
TCP/IP Stack over PCIe Bus



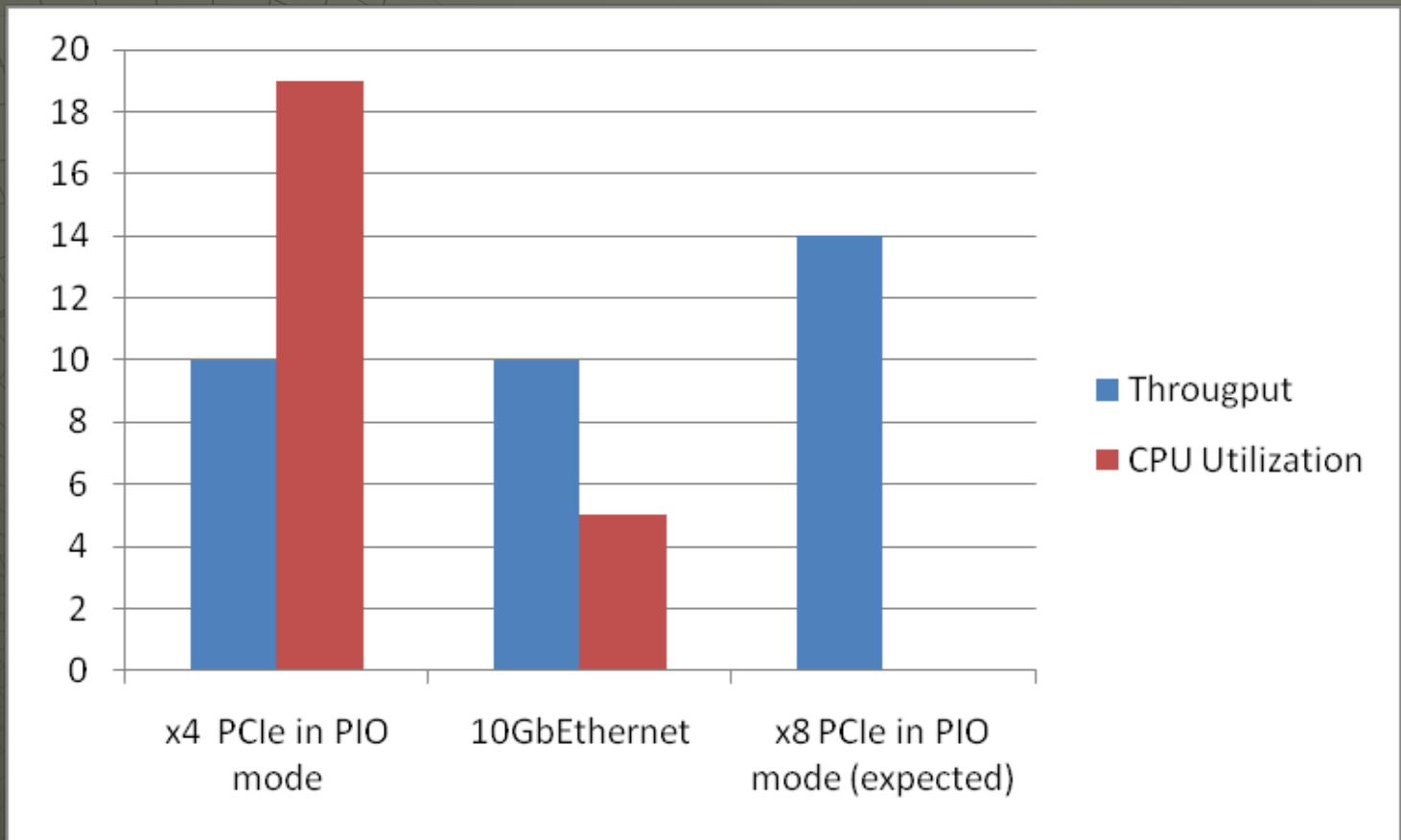
Multiple Computer Systems Connected through PCIe switch



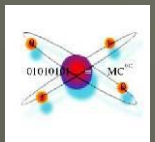
Remote system's memory is available to local system for load/store CPU instructions and DMA operations.



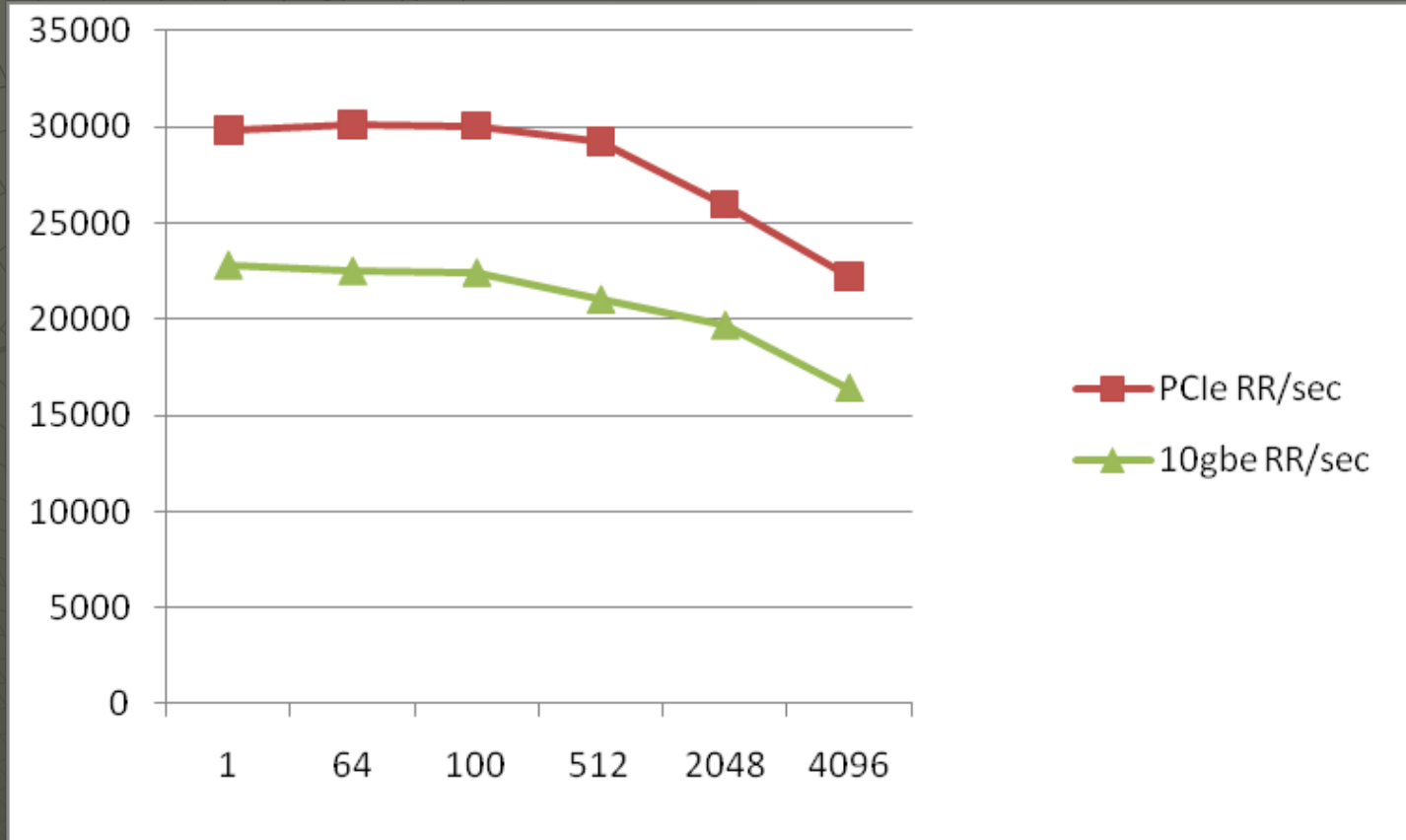
Results: Throughput & CPU Utilization



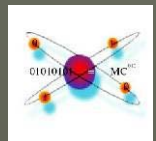
Throughput in Gbps



Results - Latency

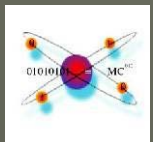


Netperf TCP_RR numbers RR/sec vs Packet size in bytes



Benefits of this Approach

- Lower Power Consumption
 - 16 lane Gen2 switch consumes 2.5W compared to 10W in NIC controller
- Lower Cost
 - Cost is around \$1 per lane, hence x8 lane device will cost \$8
- Savings in ecosystem
 - Lower cost for each component, clocks, connectors, cables, test equipments
- Lower latency and higher throughput



Issues

- Length Limitation

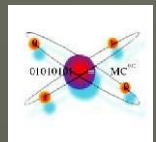
Max External cable length is approx 5m, with optical connector max length up to 100m

- Maximum number of nodes supported in a fabric

- Max number of Bus x device x func

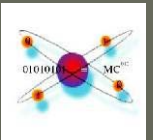
- $256 \times 32 \times 8 = 65536$ nodes

- Gen3 supports much larger number of nodes and greater cable lengths (100m) with optical



Application Areas

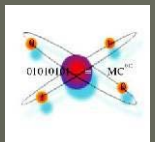
- Multi-node cluster development
- In trading systems or banking industry
- NVRAM mirroring
- Replacement of parallel buses like VME (Versa Modular Eurocard bus)



www.nucleodyne.com

- **US System Software Services company**
- Low Level Kernel & system software
- Low level device drivers for storage and communication (ethernet/PCIe) protocols
- End to end system development
 - processor, system board development
 - OS port or write new driver for OS
 - custom application development

Lets explore the possibility of supplementing your team



Two Node Communication in PIO Mode

Store r3, <address in System B's DDR>

