

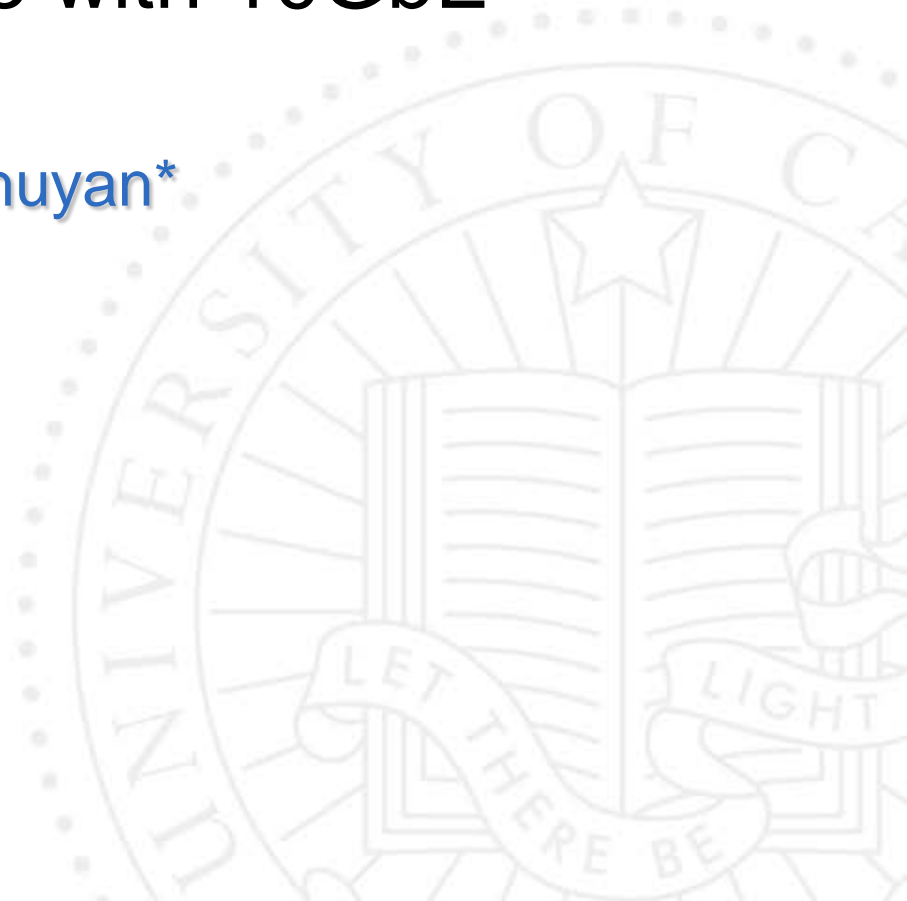
# Performance Measurement of an Integrated NIC Architecture with 10GbE

Guangdeng Liao\*, Laxmi Bhuyan\*

\* University of California, Riverside

HOT Interconnect 17, 2009

08/26/2009



# Motivation

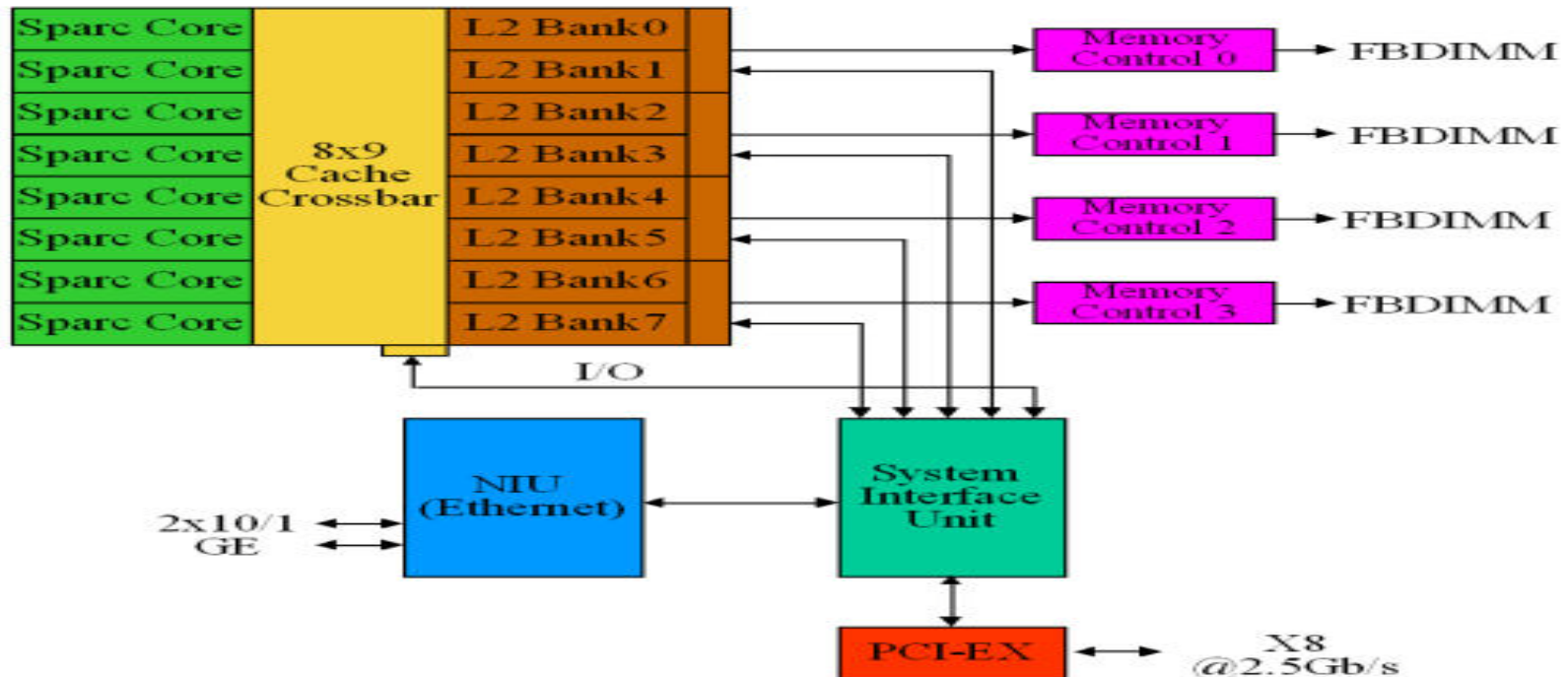
- TCP/IP Packet Processing is still a high overhead, particularly 10GbE.
- Integrated NIC based architecture has been proposed to eliminate the overhead. The performance improvement was verified by simulator.
- How an integrated NIC performs on a real machine remains unclear

# Background

- ▶ Two main overheads of TCP/IP
  - ▶ Memory Access Overhead
    - ▶ Packet Header Accesses
    - ▶ Payload copy from kernel space to user space
  - ▶ Device Driver Overhead due to long access latency to PCI-based registers
- ▶ Integrated NIC Architectures
  - ▶ Smaller Access latency to NIC registers
  - ▶ Injecting packets directly into CPU caches

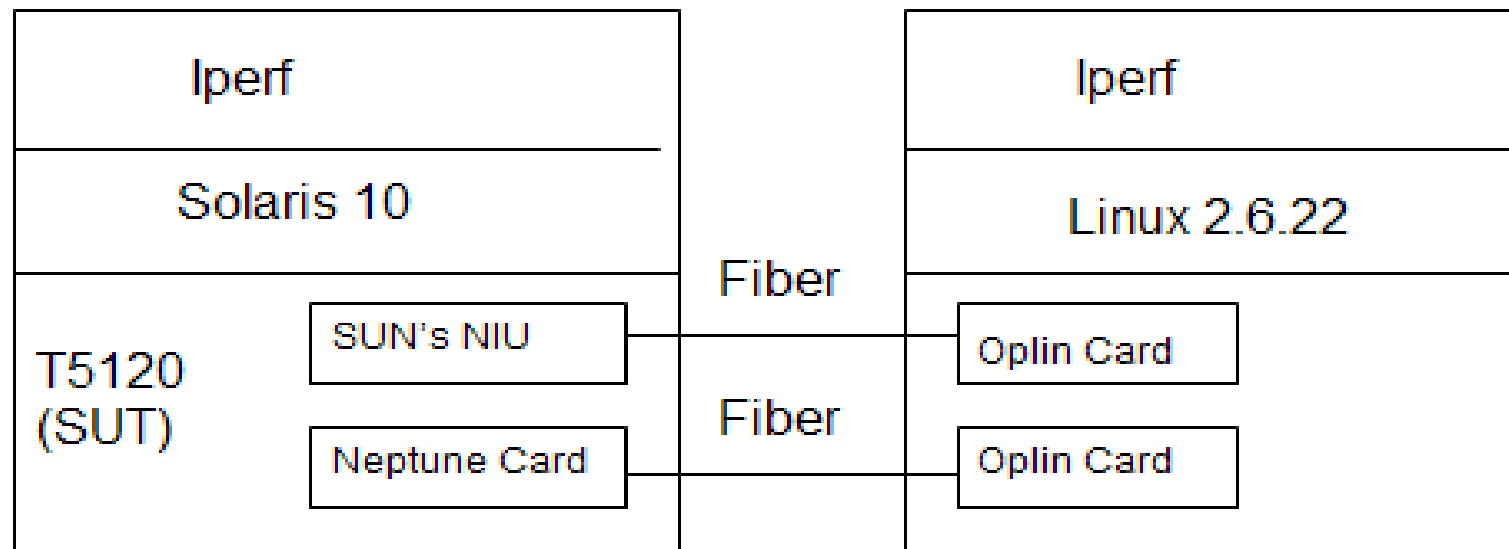
## Integrated NIC Architecture (Sun Niagara2 )

- Integrating two 10GbE NICs into CPU. Note that incoming packets are still destined to main memory instead of CPU caches.



# Experiment Methodology

- ▶ An apple-to-apple performance comparison between a PCI-based 10GbE NIC (DNIC) and an integrated 10GbE NIC (INIC)

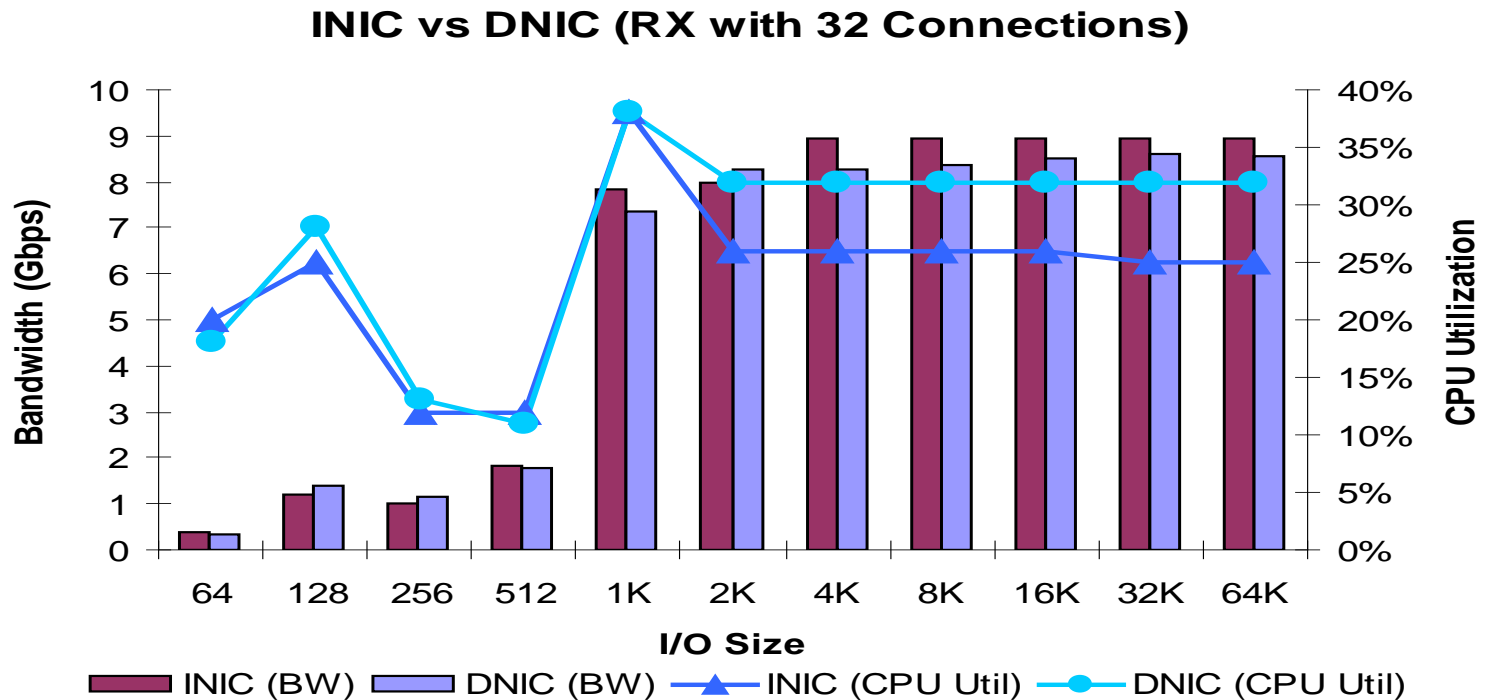


# INIC vs DNIC

Features	NIU(INIC)	Neptune (DNIC)
Transmit DMA Channels	8	12
Receive DMA Channels	8	8
Bus interface	No	8 lane PCI Express
Bus bandwidth limit	No	16 Gbits/s each direction
Transmit Packet Classification	Software	Software
Receive Packet Classification	Hardware	Hardware

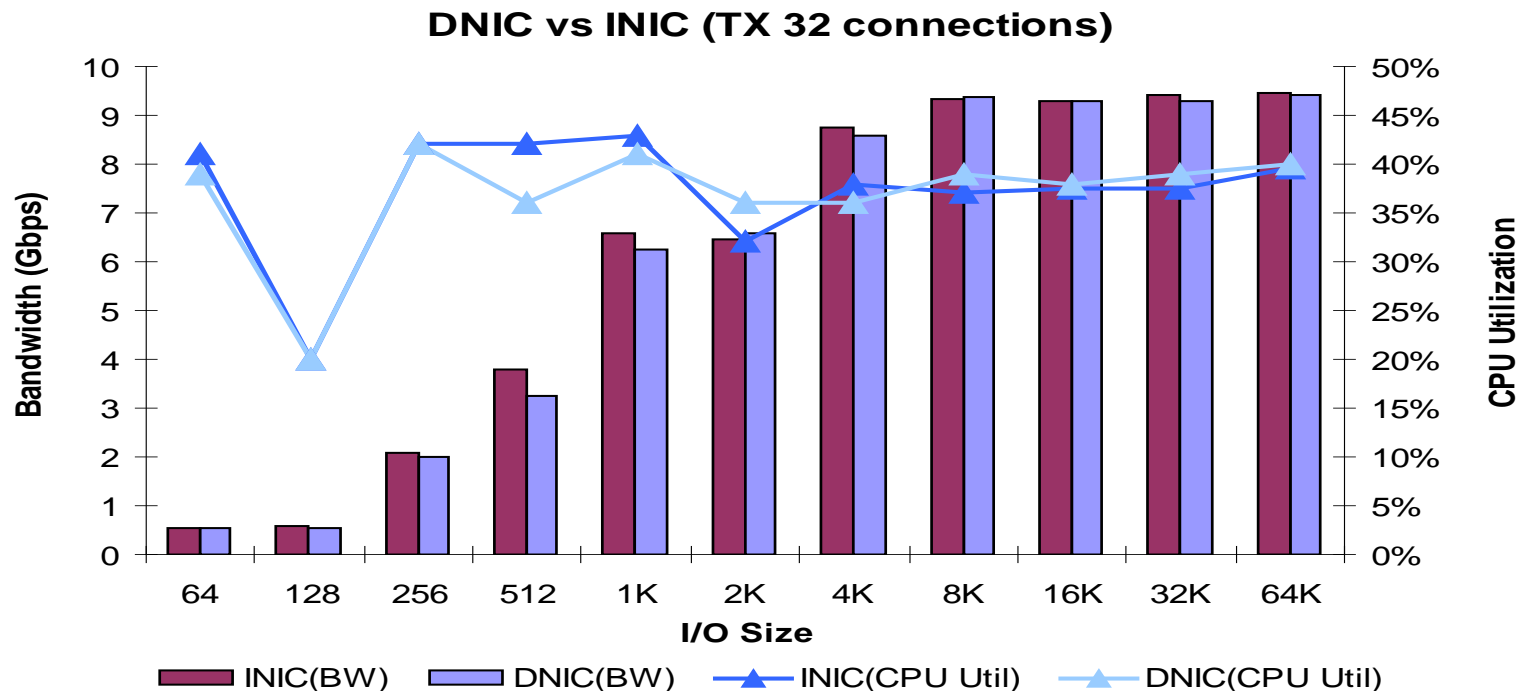
# Performance Comparison with 10GE (RX)

- › The INIC obtains 7.5% higher bandwidth and saves 20% relative CPU utilization on average for large I/O sizes (>1KB).



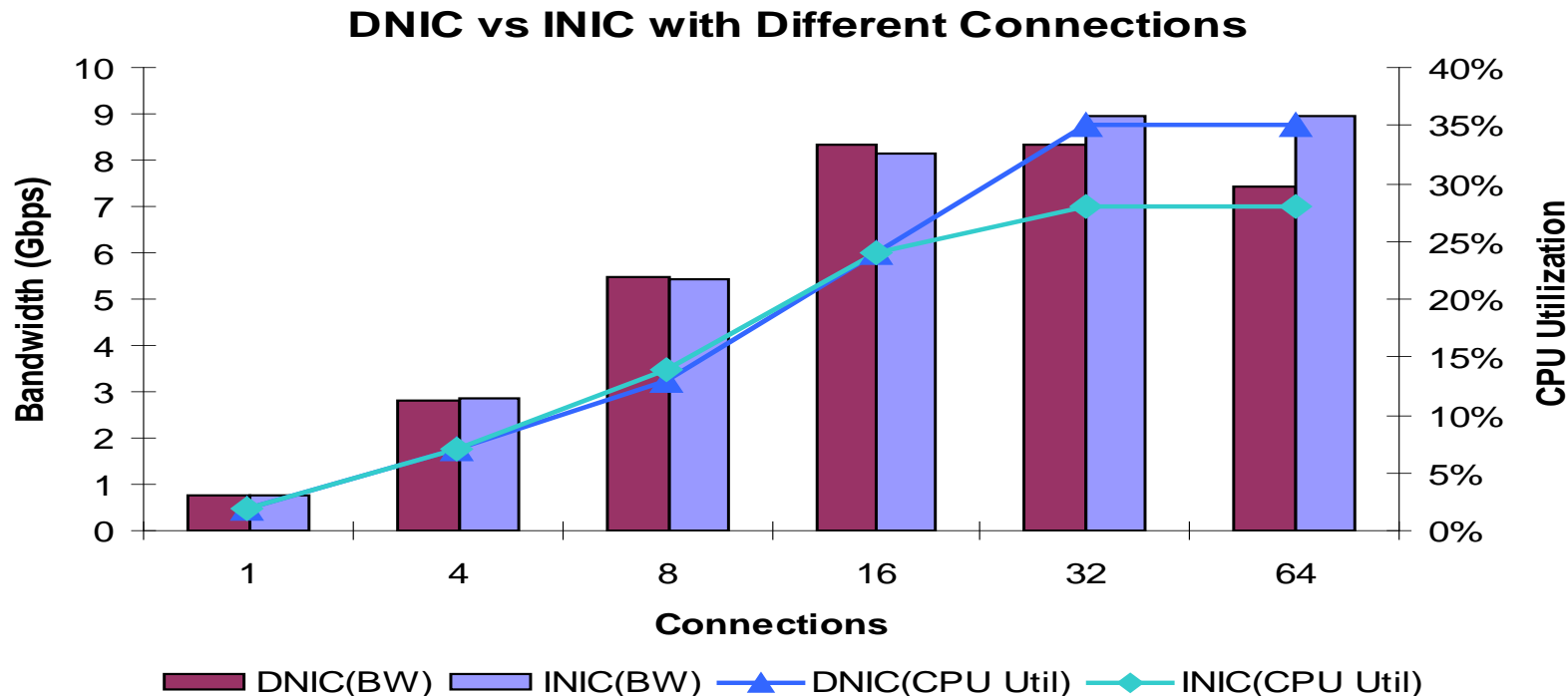
# Performance Comparison with 10GbE (TX)

- Unlike the RX side, INIC and DNIC have the similar transmit network performance.



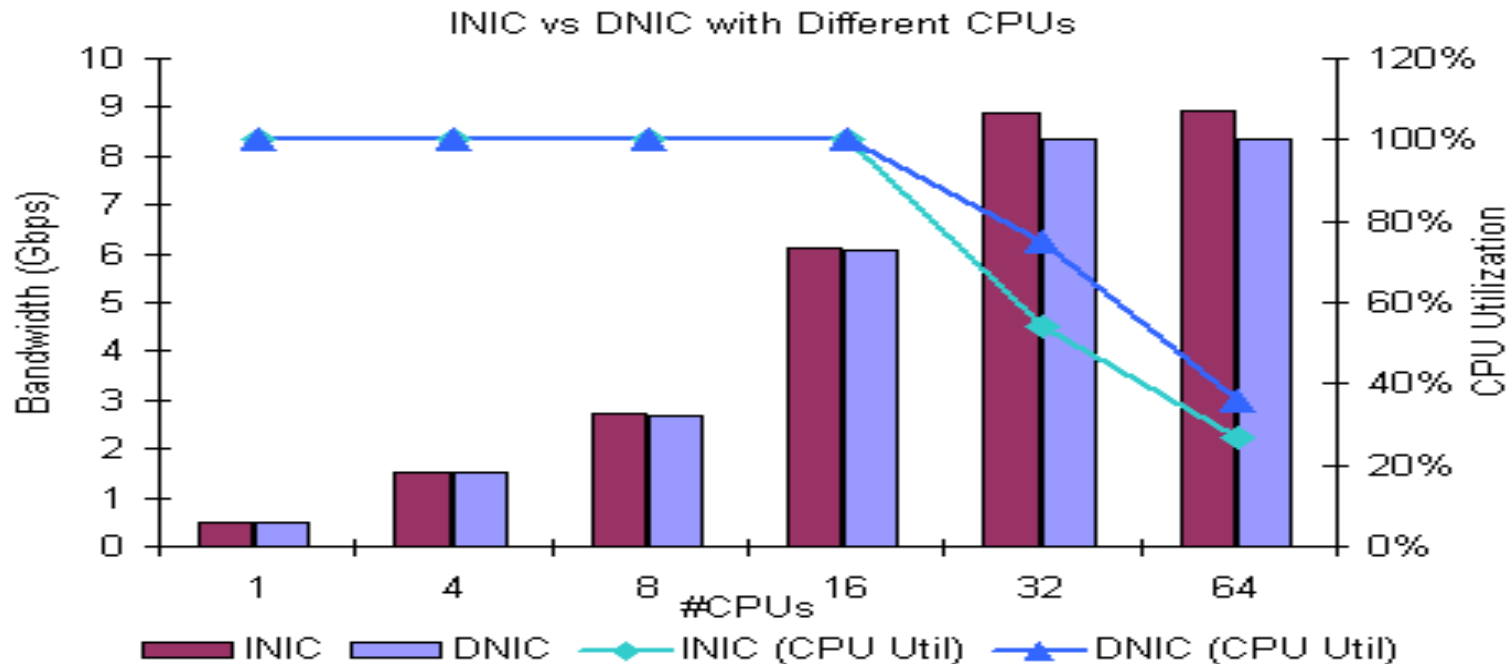
# Performance Comparison (RX) with Various Connections

- › INIC performs better than DNIC only with greater than 16 Connections.



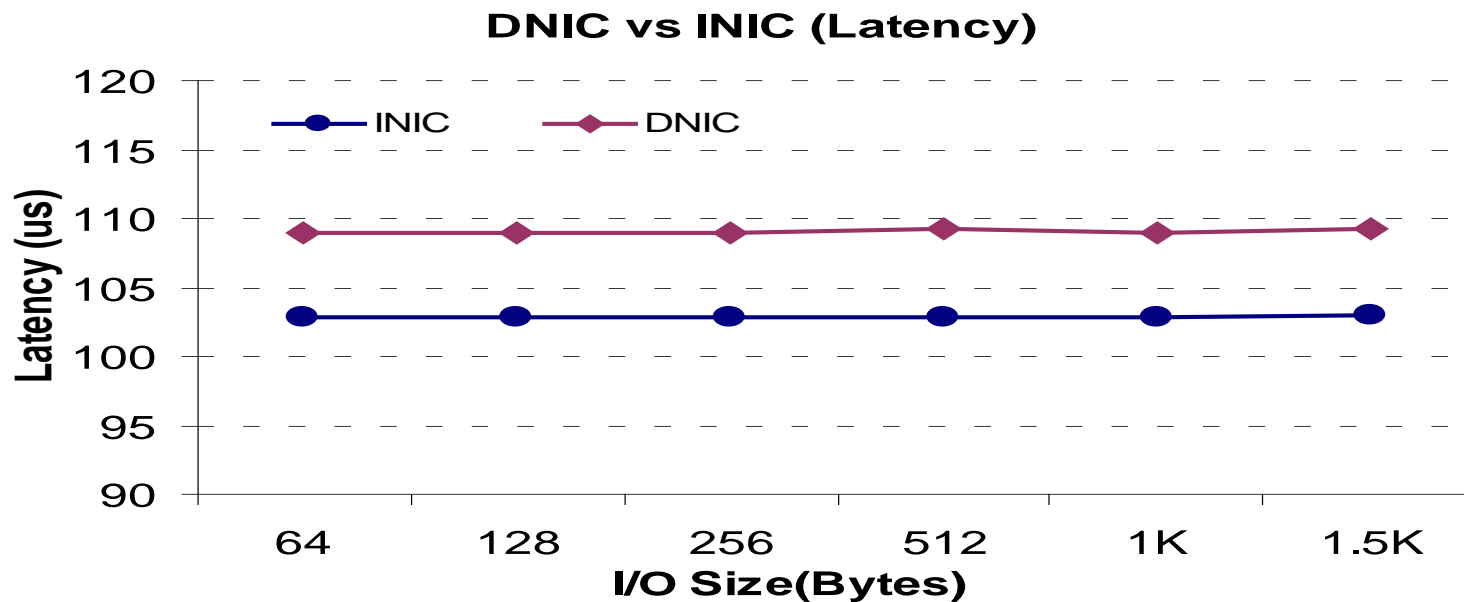
## Performance Comparison (RX) with Various CPUs

- INIC performs better than DNIC only with more than 16 CPUs.



# Latency Comparison

INIC can achieve a lower latency by saving 6  $\mu\text{s}$ . It is due to the smaller latency of accessing I/O registers and eliminating PCI-E bus latency.

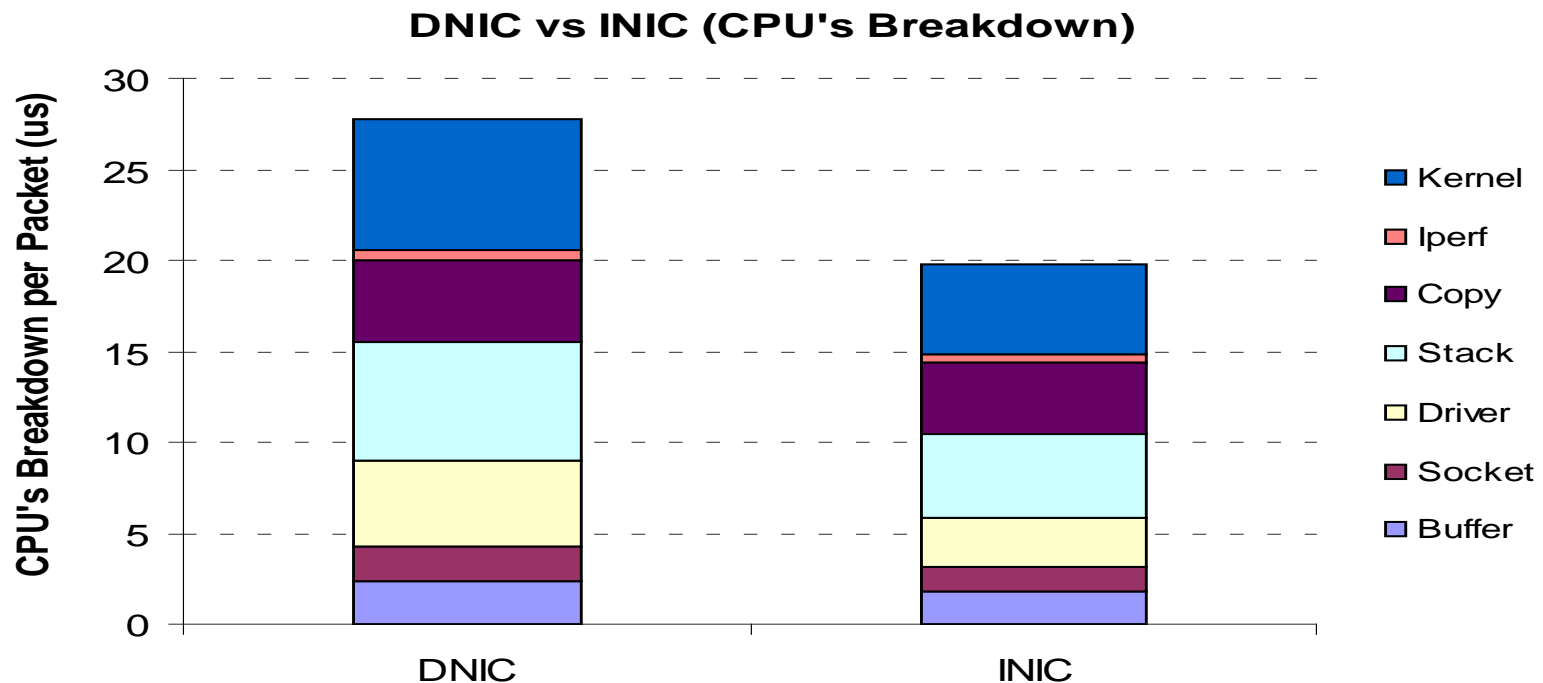


# Observations

- › INIC improves RX network efficiency only with greater than or equal to 32 connections.
- › the integration could affect the system behaviors with a large number of connections, and different system behaviors mainly cause the performance difference.
- › the benefits can only be achieved with large number of CPUs, and thus are tied to the highly threaded Sun system.

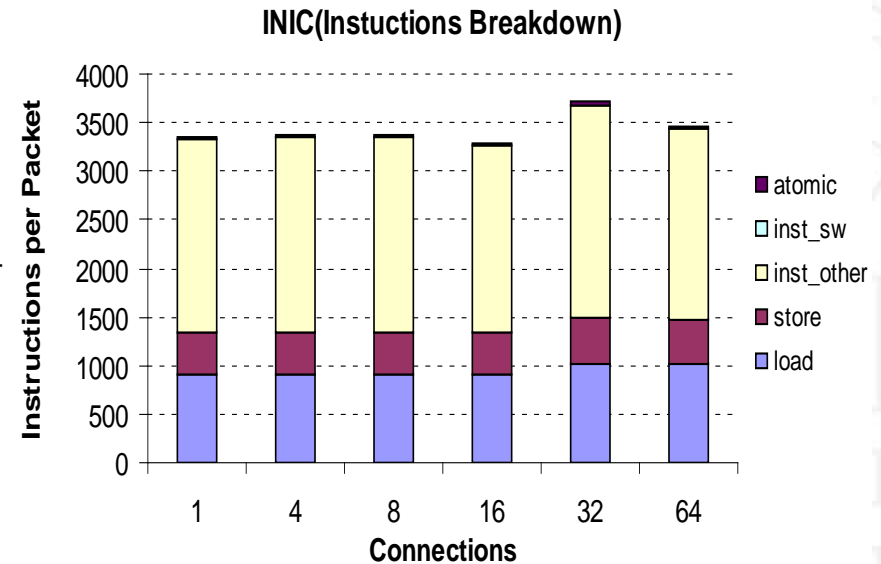
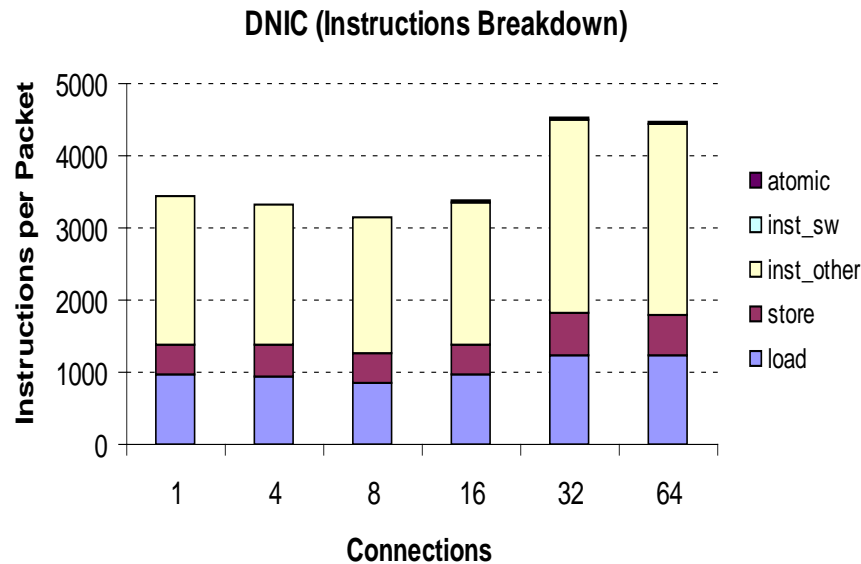
# Performance Characterization

- Per-packet processing overhead breakdown shows that INIC reduces the driver overhead by half. Surprisingly, INIC also reduces the overheads on network stack, buffer management, socket and kernel, and they mainly contribute to the performance benefits.



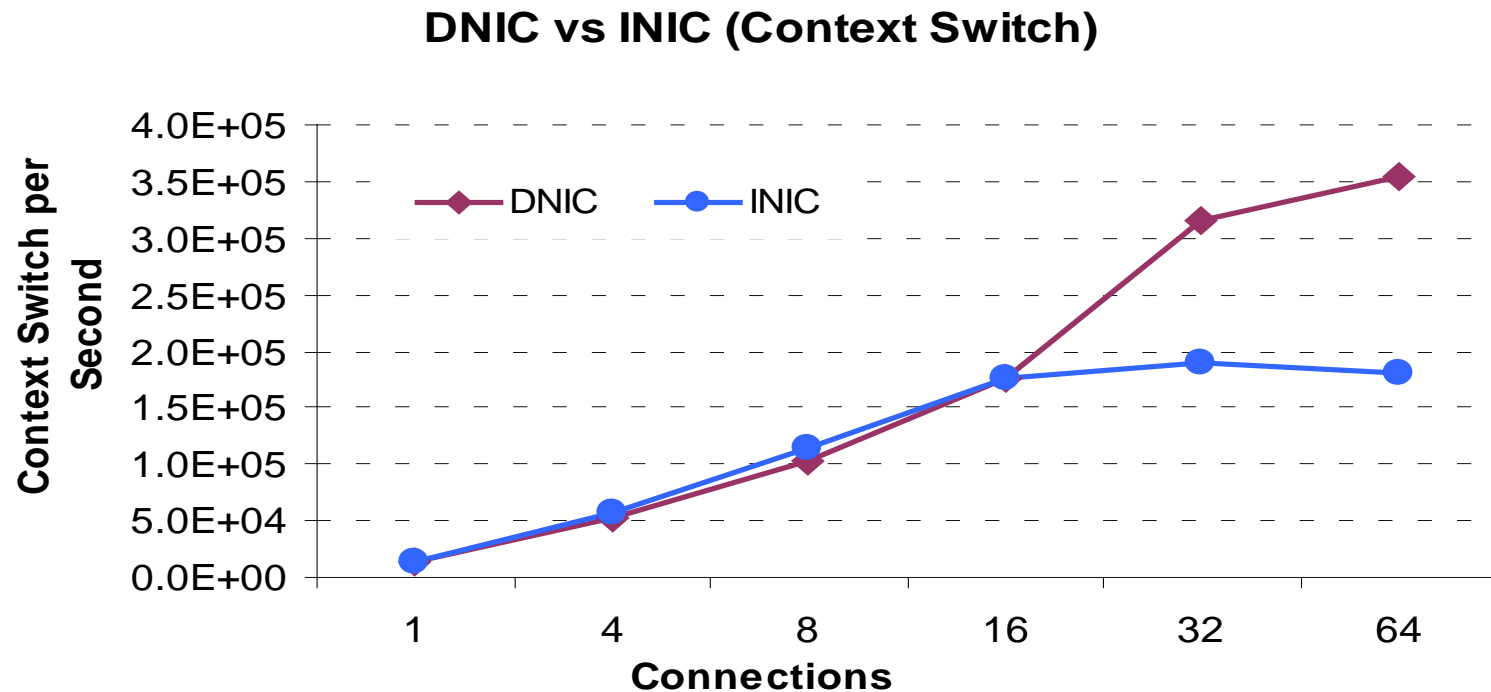
# Per-packet Instruction Breakdowns

- With DNIC, per-packet instructions increase when the number of connections is greater than 16. Contrary to DNIC, increased connections with INIC do not significantly increase instructions per packet.



# Context Switch Rate

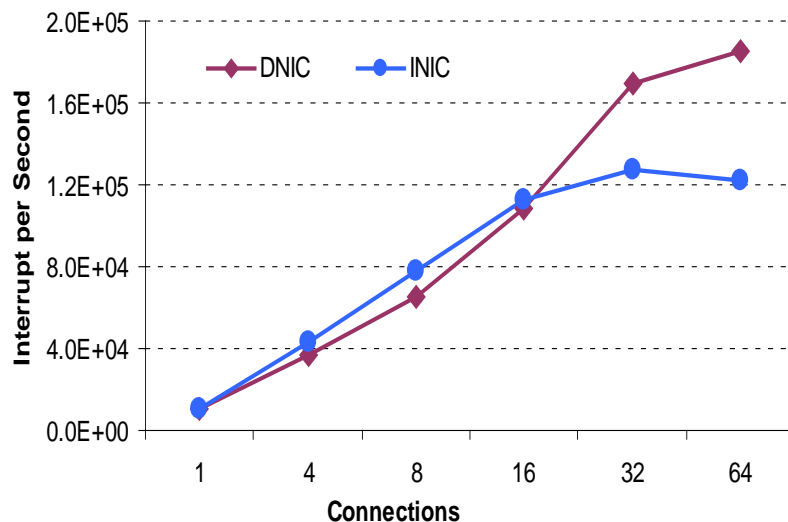
- Higher context switch rate with DNIC explains the higher numbers of per-packet load/store instructions with DNIC.



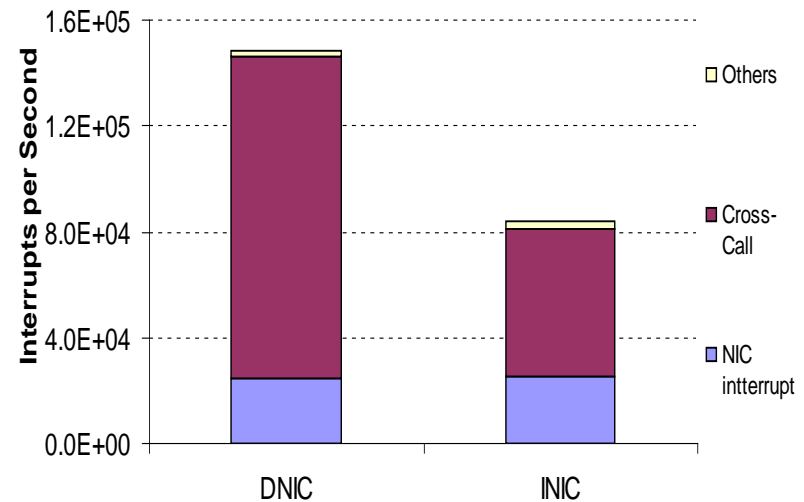
# Interrupt Rate

- › DNIC has higher interrupt rate than INIC when the number of connections is greater than 16, and the interrupt rate gap is mainly from cross-calls or inter-processor interrupts.
- › Higher Interrupt rate with DNIC explains the higher context switch rate with DNIC.

DNIC vs INIC (Interrupt per Second)

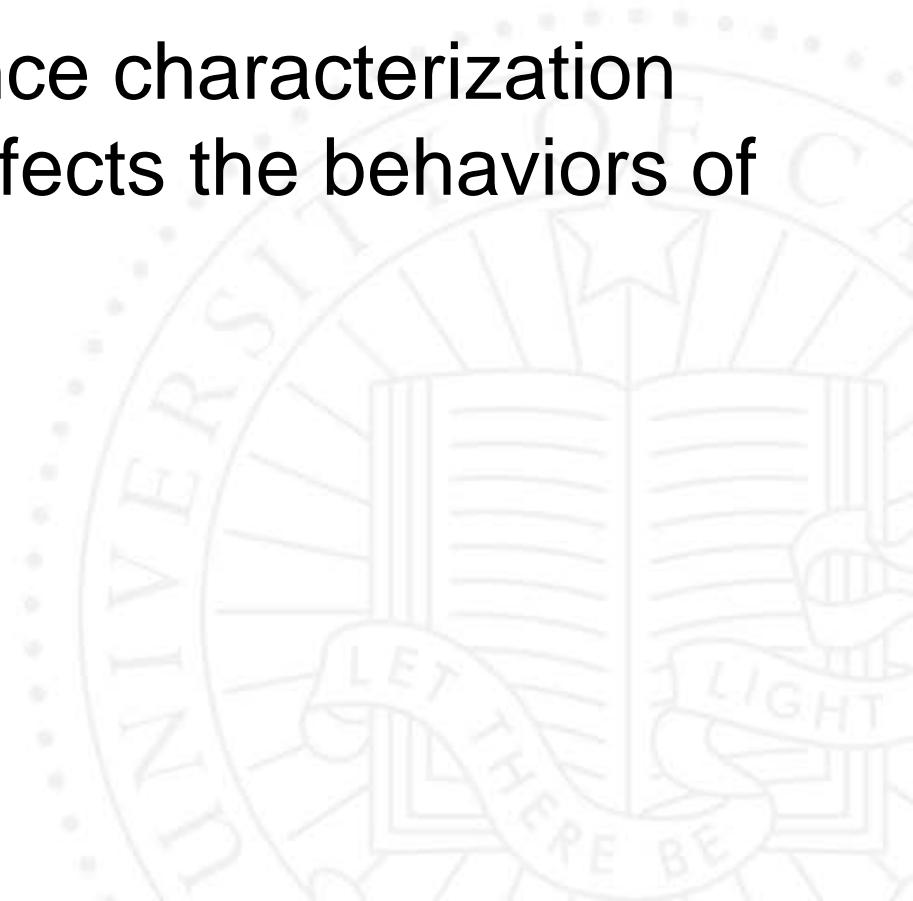


DNIC vs INIC (Interrupt Breakdown)



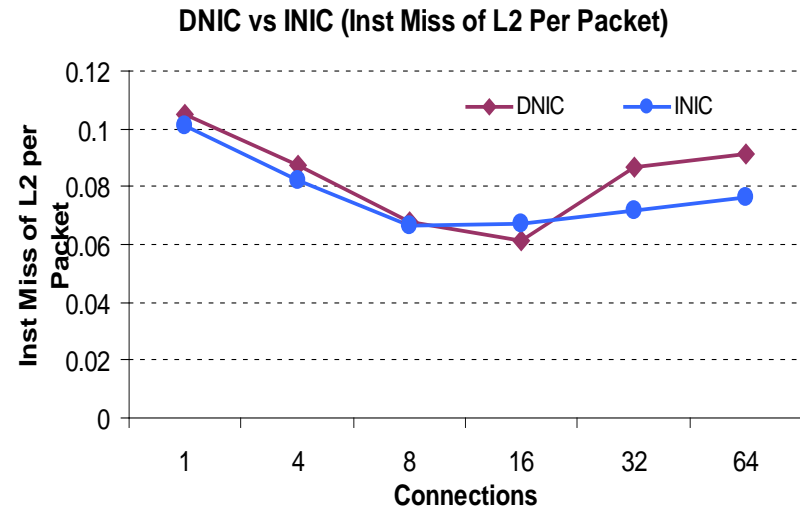
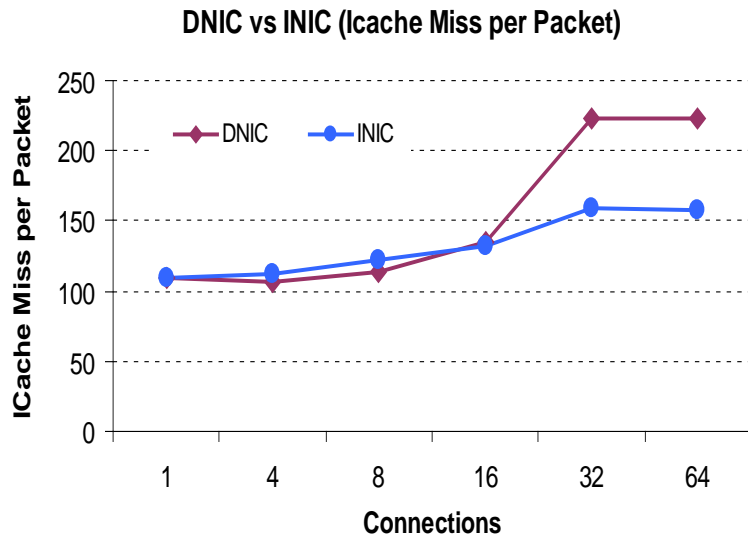
# Observation1

- The detailed performance characterization reveals that the INIC affects the behaviors of the OS scheduler.



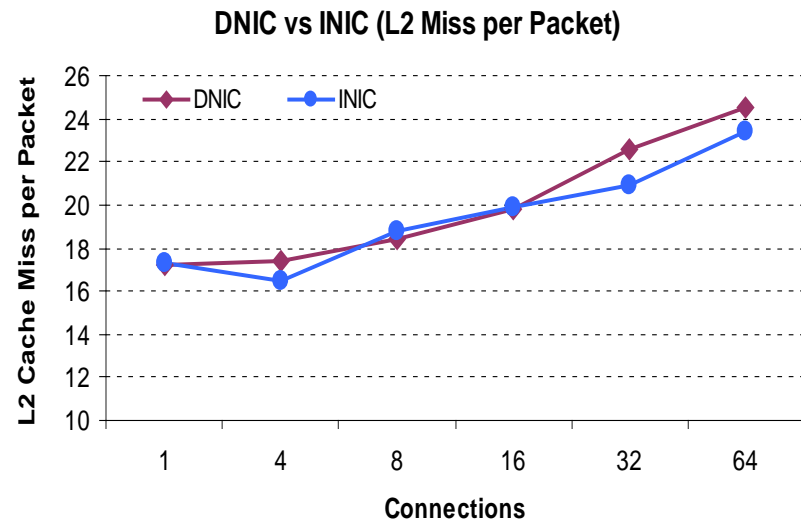
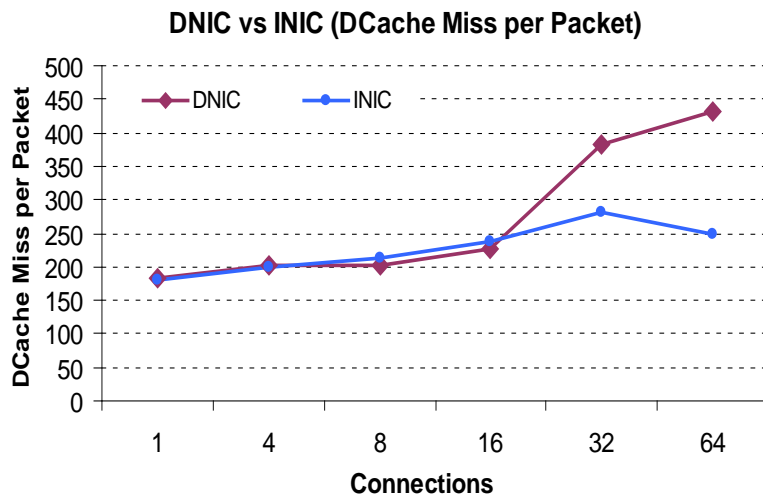
# Instruction Cache and L2 Cache

- More instruction cache misses per packet happen with DNIC when the number of connections is greater than 16.
- Instruction misses rarely happens for both DNIC and INIC in the shared L2 cache.



# Data Cache and L2 Cache

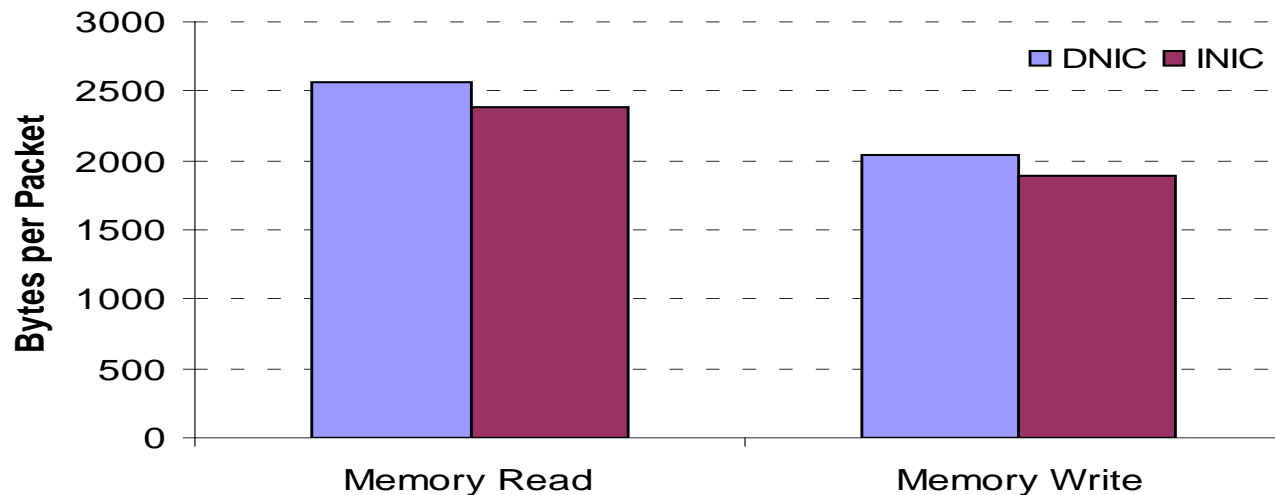
- More data cache misses per packet happen with DNIC when the number of connections is greater than 16.
- In L2 cache, the INIC has 7.6% reduction of data misses or 1.5 less data misses.



# Memory Access Traffic

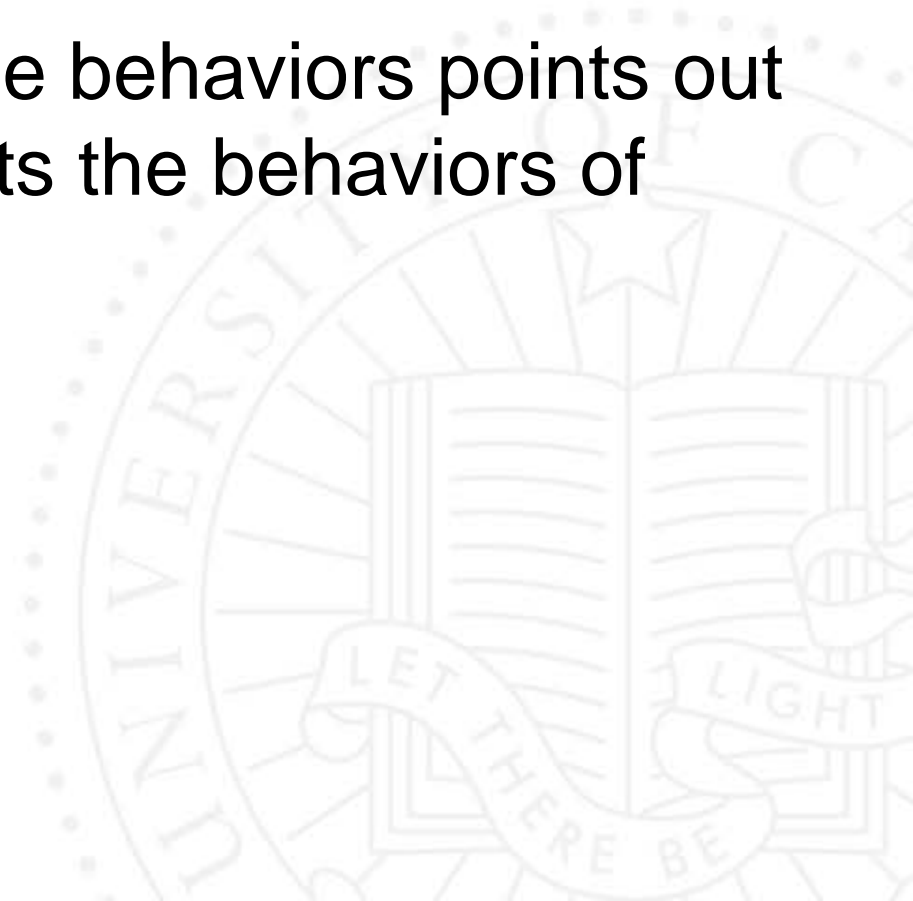
- More cache misses reported before lead to more memory accesses and thus increase memory access traffic.

**DNIC vs INIC (Memory Traffic per Packet)**



# Observation 2

- ▶ The study of CPU cache behaviors points out that the INIC also affects the behaviors of CPU caches.



# Key Takeaways

- › Our measurement on the real machine confirms the observation from the simulator that the driver overhead is significantly reduced by integrating NICs into CPUs.
- › Surprisingly, overheads in other components are also reduced and their improvements mainly contribute to the performance benefits achieved by INIC.
- › Our detailed characterization shows that the INIC affects the behaviors of both the OS scheduler and CPU caches.
- › Those performance benefits achieved by INIC is tied to the tested highly threaded Niagara2 processor.
- › A new interaction mechanism between CPUs and NICs is needed in the next generation I/O infrastructure.



# Q & A

