

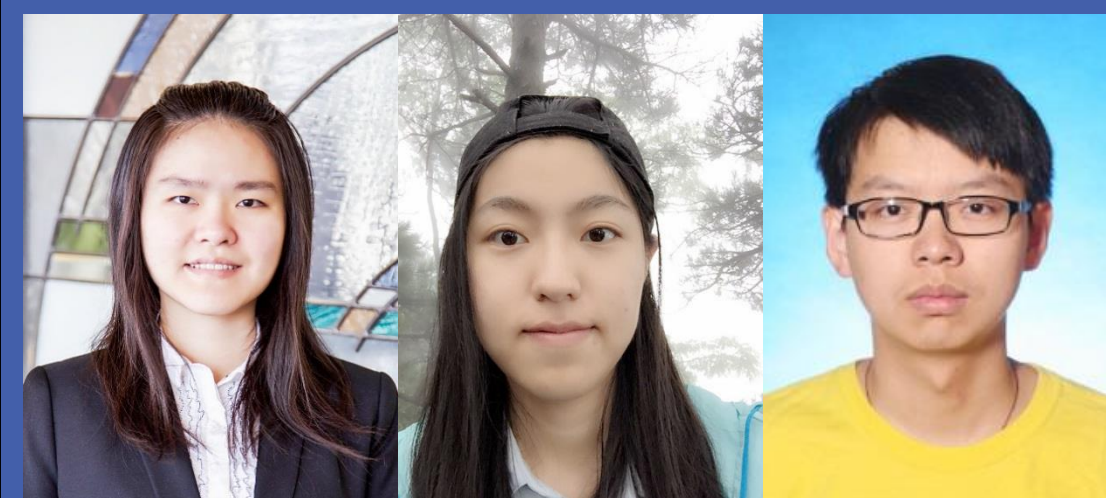
Network Interface Buffer Elimination



EECS578 Fall'15

Qilu Guo, Jing Ji, Jiong Xue

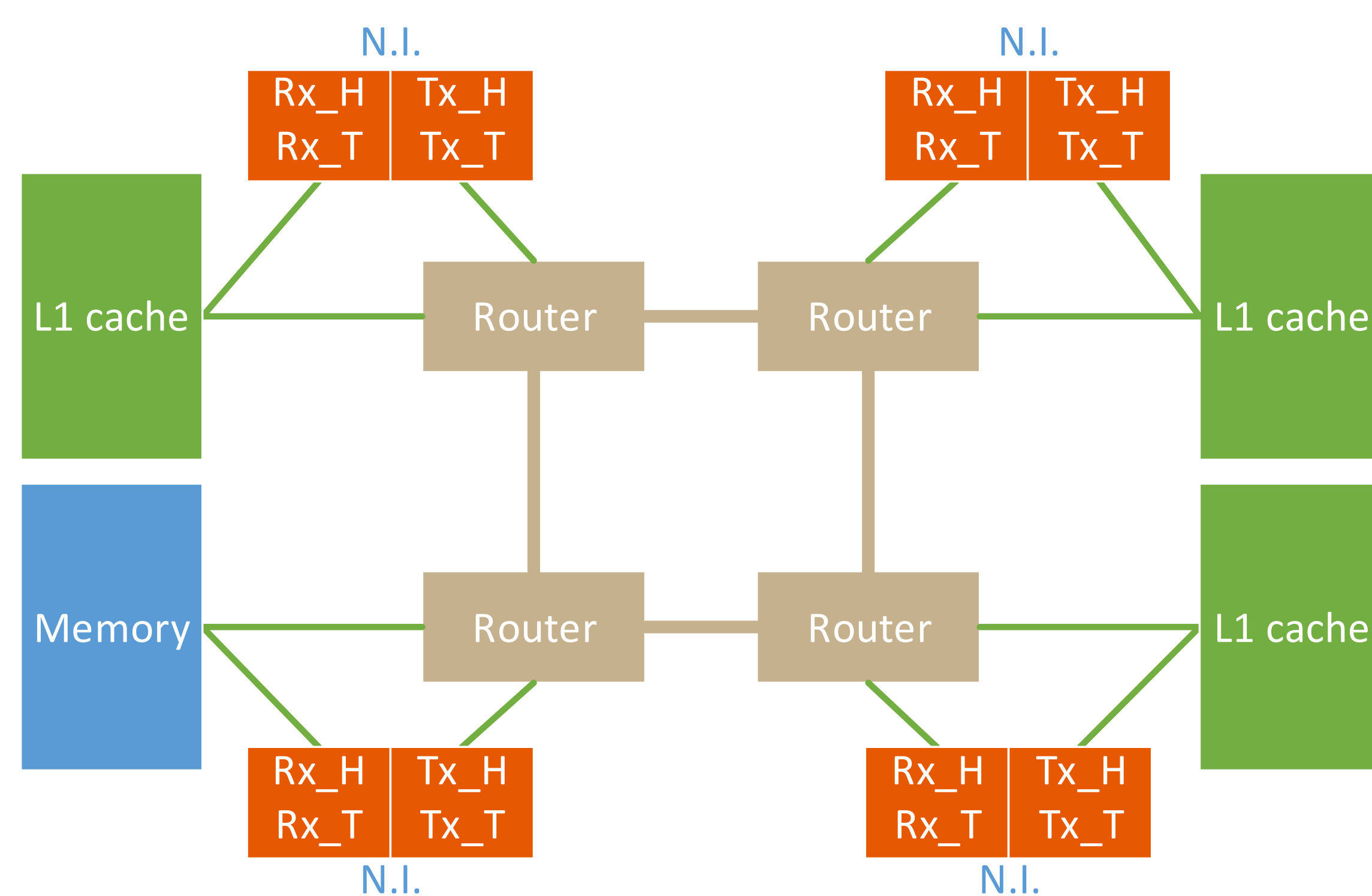
{qiluguo, silverji, xuejiong} @umich.edu



Motivation

- Multiprocessor architectures and platforms have been introduced due to more aggressive transistor scaling and larger performance gap between Moore's Law and the art-of-design technology
- Network-on-chip is a general purpose communication concept that provides high throughput, but comes with high expense in both area and power
- Reducing area brings lower power consumption and less cost
- Network interface plays significant role in determining overall network-on-chip area

Proposed Solution



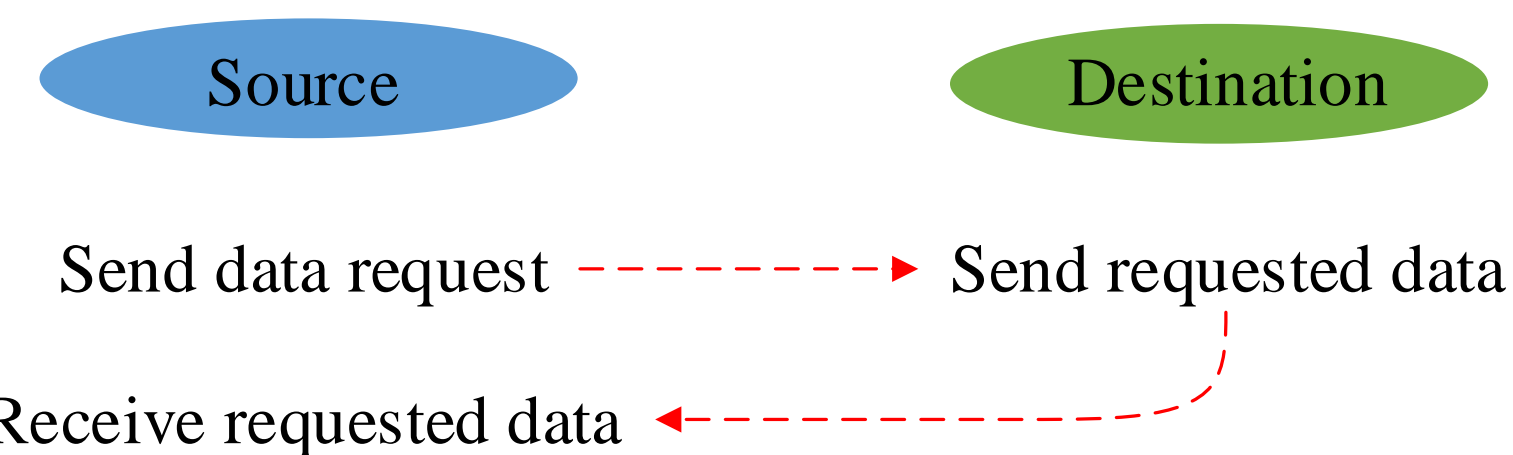
- The network interface only stores the head and the tail flits and the data are preserved in cache
- Data are allowed to transmit between the cache and the router directly

Comparison

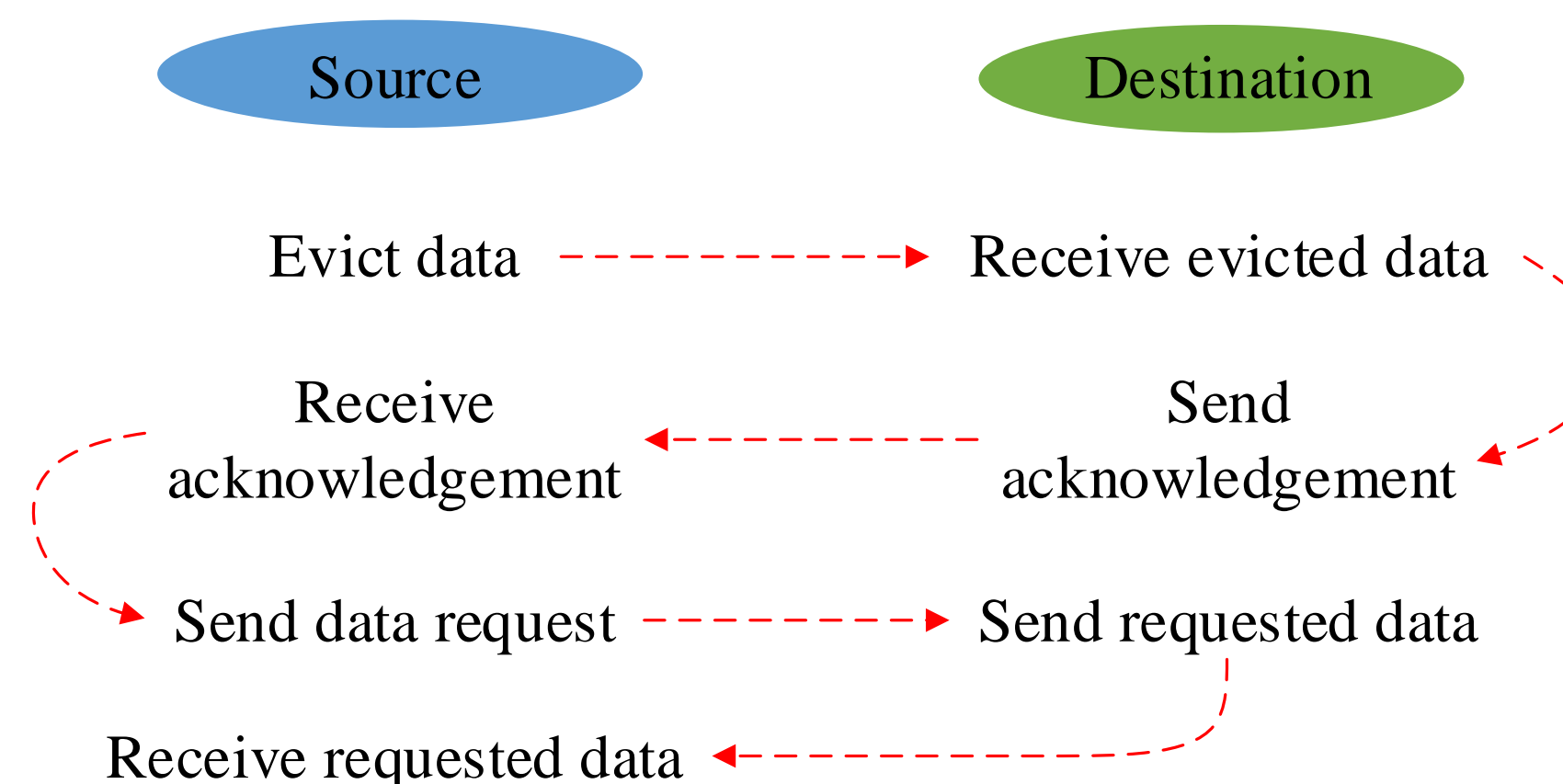
Components	Baseline	Our Design
Cache	N-way write-back, write-allocate cache Blocked until the transmission completes	
Network Interface	Store head, data and tail flits	Only store head and tail flits
Router	Communicate only with the network interface	Communicate with both the network interface and the cache

Data Transmission Protocol

- When a cache miss happens on an invalid cache line, the cache will send an request to the processor to ask for the data from other caches.

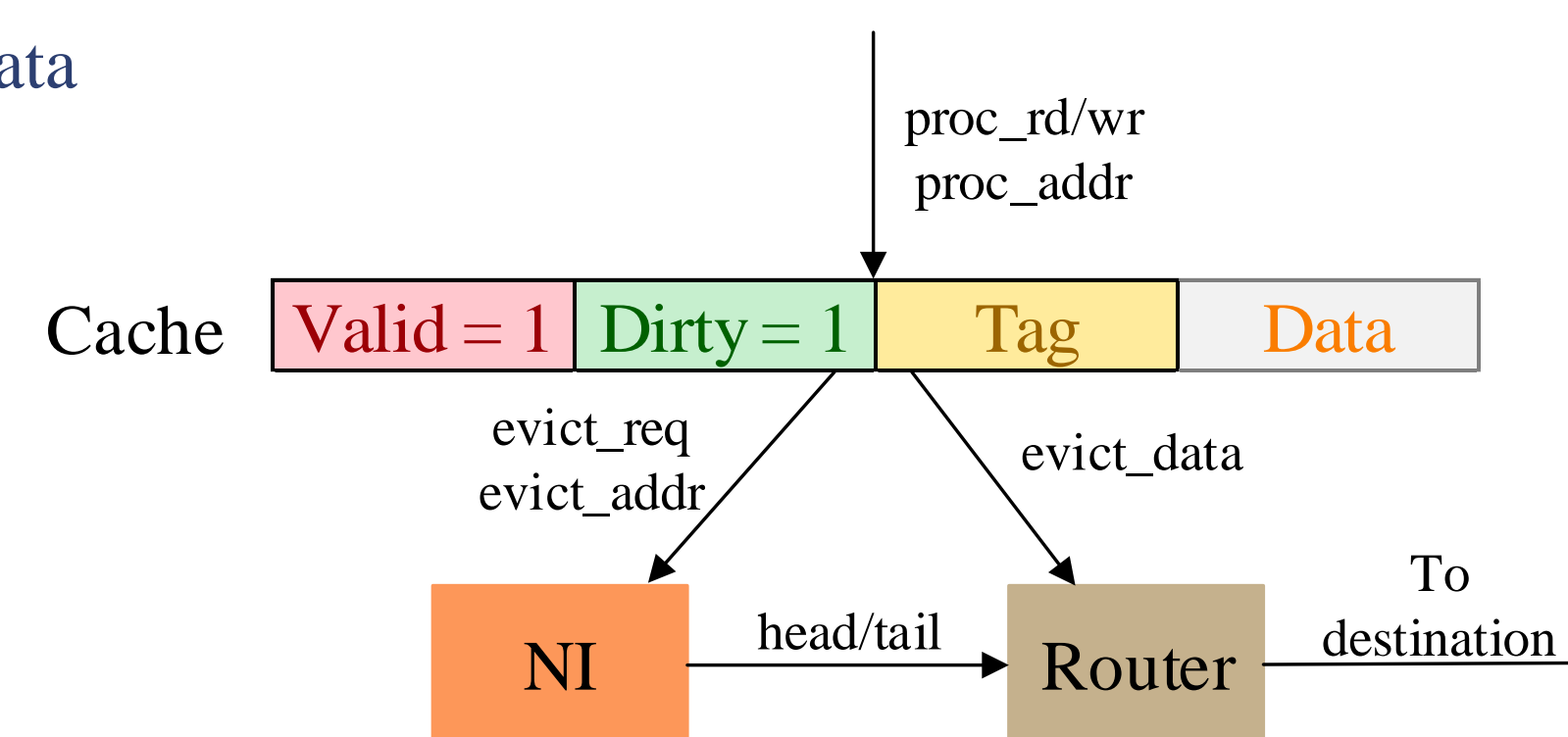


- When a cache miss happens on a valid cache line, eviction of that cache line is needed first.

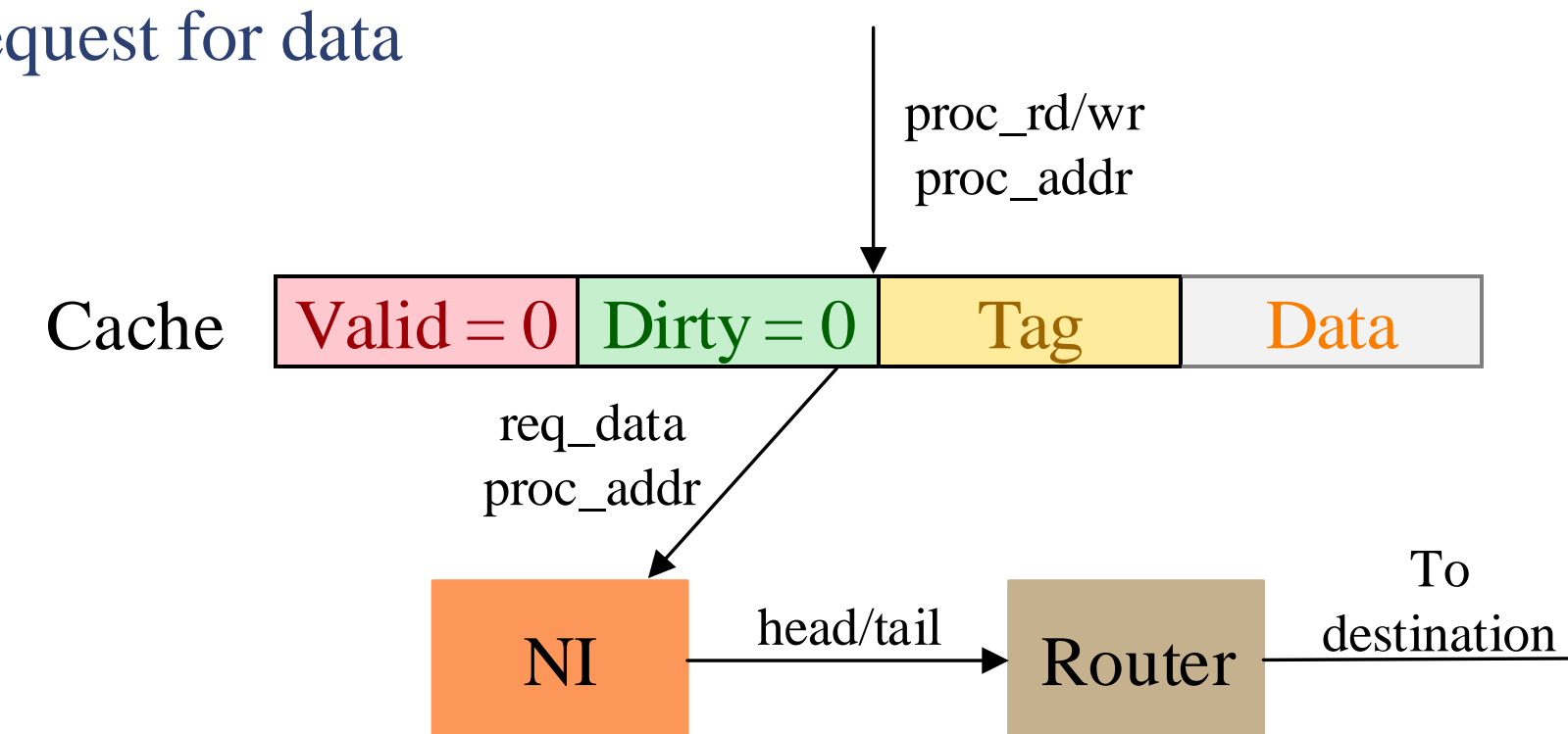


- Three typical types of data flow

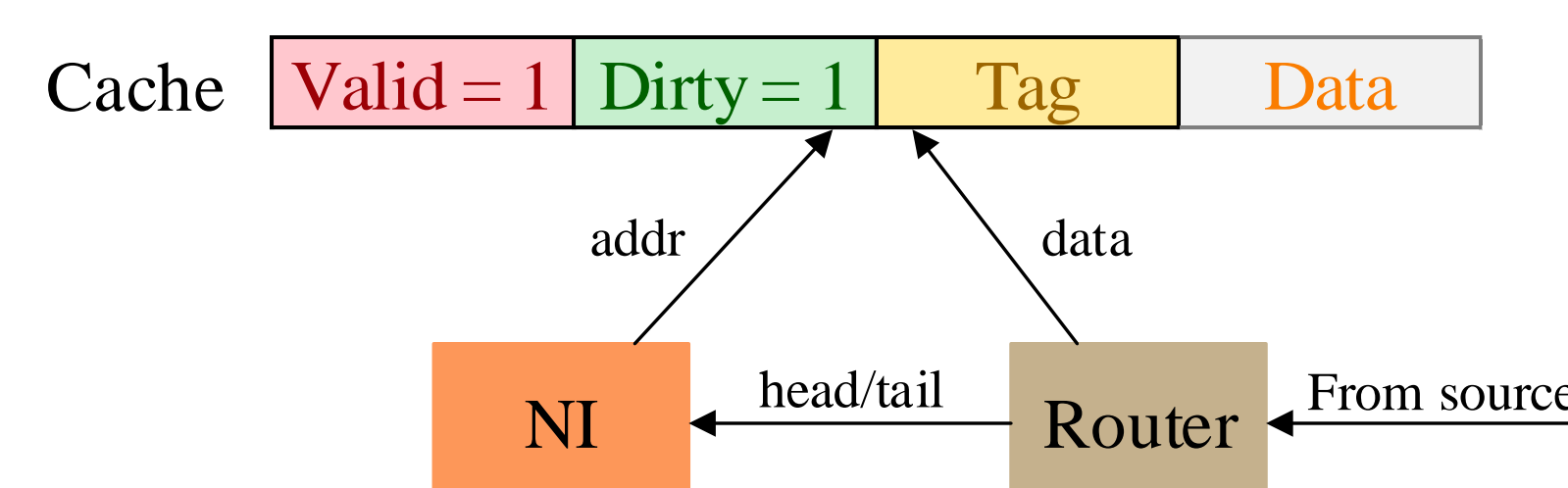
- Evict data



- Send request for data



- Receive data



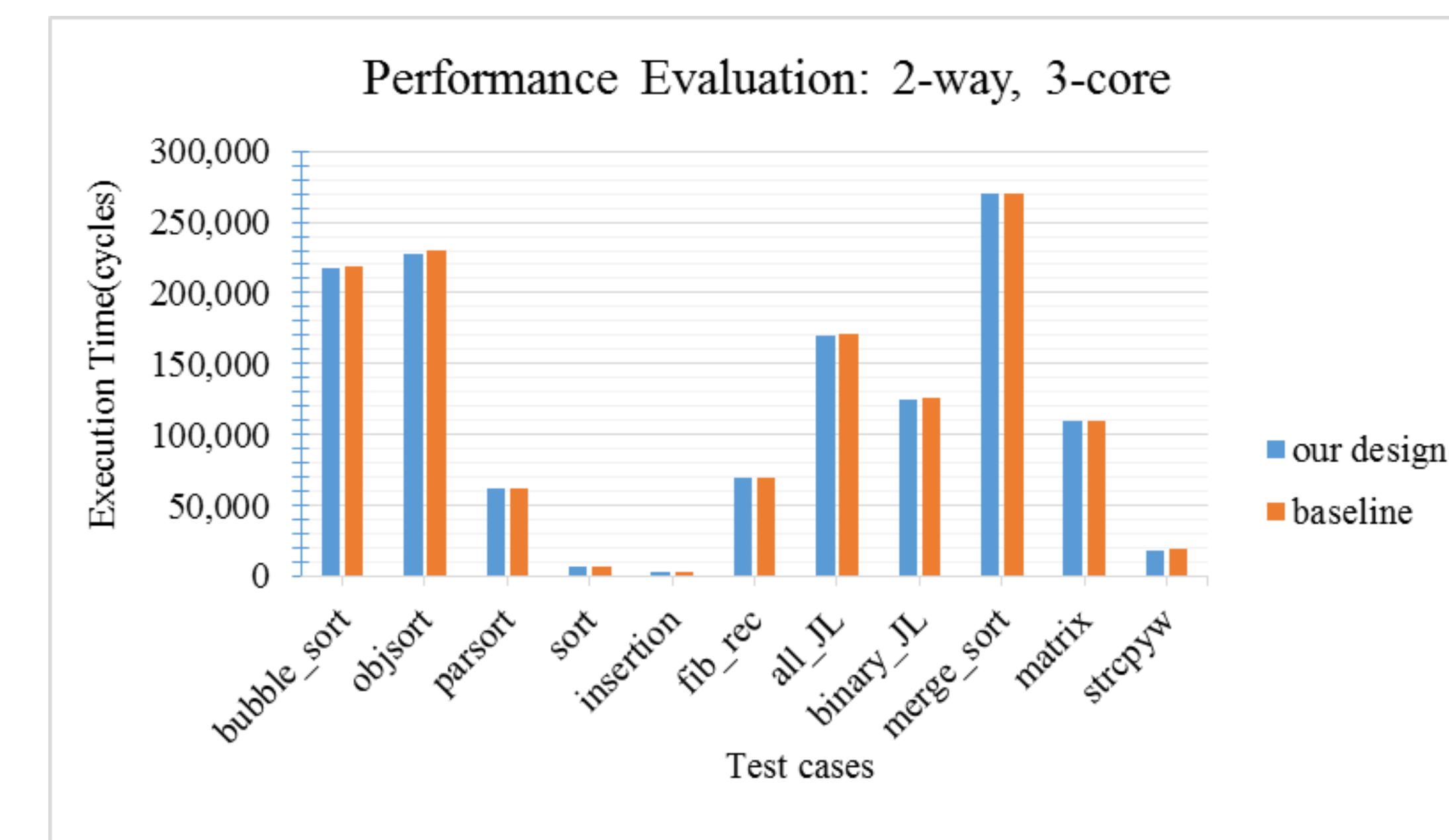
Experimental Evaluation

Area

module	baseline (μm^2)	our design (μm^2)
router	512,292	518,037
network interface	179,558	31,518

- Synthesized the RTL codes by Synopsys Design Compiler
- The cache module is unmodified, so the area remains the same
- The area of the router increases by 1.1% due to the extra ports and control logic to communicate with the cache
- The area of the network interface is reduced by 6 times resulting from the buffer elimination

Performance



- We obtained memory access traces by running test cases on the EECS 470 processor
- We measured execution cycles by injecting traces to the caches
- Performance improvement ranges from 0.43% - 1.32%, with the average being 0.77%
- Performance improvement results from direct injection of data flits to the router without going to the network interface

Conclusion

- Our design successfully reduces the area of the network interface by 6 times, while keeping the cache the same and having an extremely low area overhead for router
- Our design does not degrade the performance but instead improves the performance with the average of 0.77%