Network Balancing Act
A High-performance Packet Processing Framework for Heterogeneous Processors

Joongi Kim, Keon Jang, Keunhong Lee, Sangwook Ma, Junhyun Shim, and Sue Moon
Department of Computer Science, KAIST, (joongi, keonjang, khlee, sangwook, junhyun)@an.kaist.ac.kr, sbmoon@kaist.edu

Challenges of SW Packet Processing

- High-performance + Programmability
  - Processed by NUMA multi-core CPUs
  - 10 GbE NICs

- Element interface & HW resource mapping
  - User-defined function

- Generalized model for the load balancer
  \[ E = \frac{1}{\text{Throughput}} = w \times \text{PPC}_{\text{GPU}} + (1 - w) \times \text{PPC}_{\text{CPU}} \]
  ※ PPC: amortized per-packet processing cycles, w: offloading weight

- Extending Work

Our Contributions

- Batch-oriented packet processing pipeline
  - Separate system resource mapping to avoid branches
  - Branch prediction to make branches faster

- Offloading abstraction
  - Declarative I/O data formats \( \Rightarrow \) Datablocks

- Adaptive CPU/GPU load balancing
  - Optimize toward maximum throughput

- Multi-core & NUMA scalable implementation
  - Basing on Intel DPDK (data-plane development kit)
  - Scripting interface to map system resources

Resulting Performance

- Adaptive load balancer: over 90% of throughput compared to manual optimization

- Full implementation of “datablocks”
  - Left as future work in the paper
  - Currently under performance optimization (10-30% slower than the paper version)

This work was supported by National Research Foundation of Korea (Project No. 2014007580).