DPDK
SUMMIT
APAC • 2021

Handling Elephant Flow on a DPDK-Based Load Balancer

CHENMIN SUN, YIPENG WANG, HONGJUN NI
INTEL

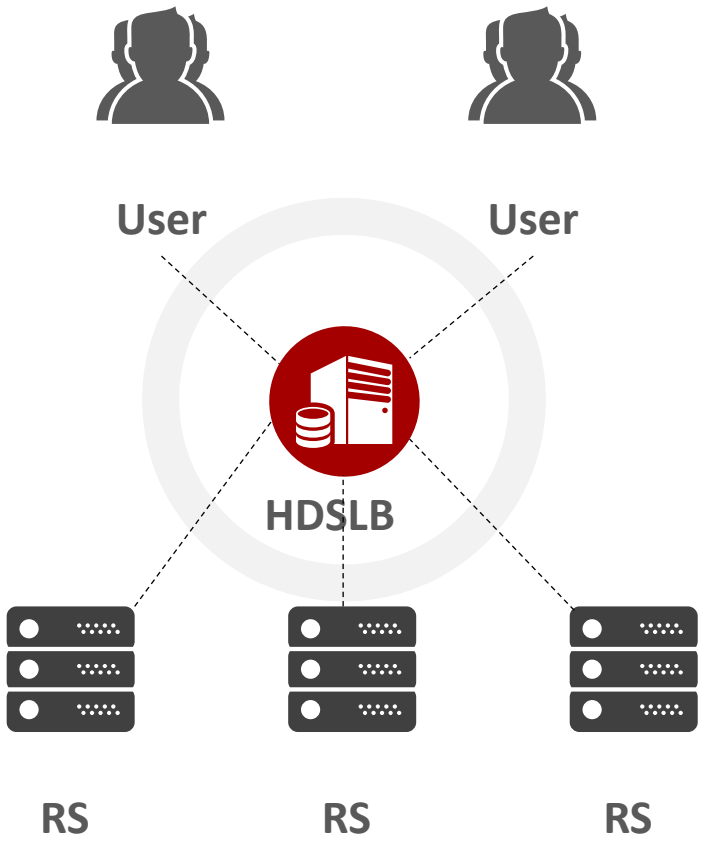
Agenda

DPDK SUMMIT APAC•2021

- HDSLB Introduction
- Problem Statement
- Innovative Algorithm
- DLB-assisted Distribution
- Key Takeaway
- Q & A

HDSLB Introduction

HDSLB: High Density Scalable Load Balancer



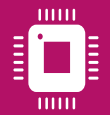
Performance Requirements for Single Node

01	100M Level Concurrent Conn
02	150Mpps/200Gbps Throughput
03	Single Session 10Mpps Level

HDSLB Highlights

DPDK SUMMIT APAC•2021

HDSLB Addressing These Challenges With Industry Leading Performance



Intel Processors and NIC Packaged Solution

Fully optimized

Handle 100M Level Concurrent Conn

Address the business challenges for large concurrent conns

Handle 150Mpps Level Throughput

Address the business challenge of huge traffic

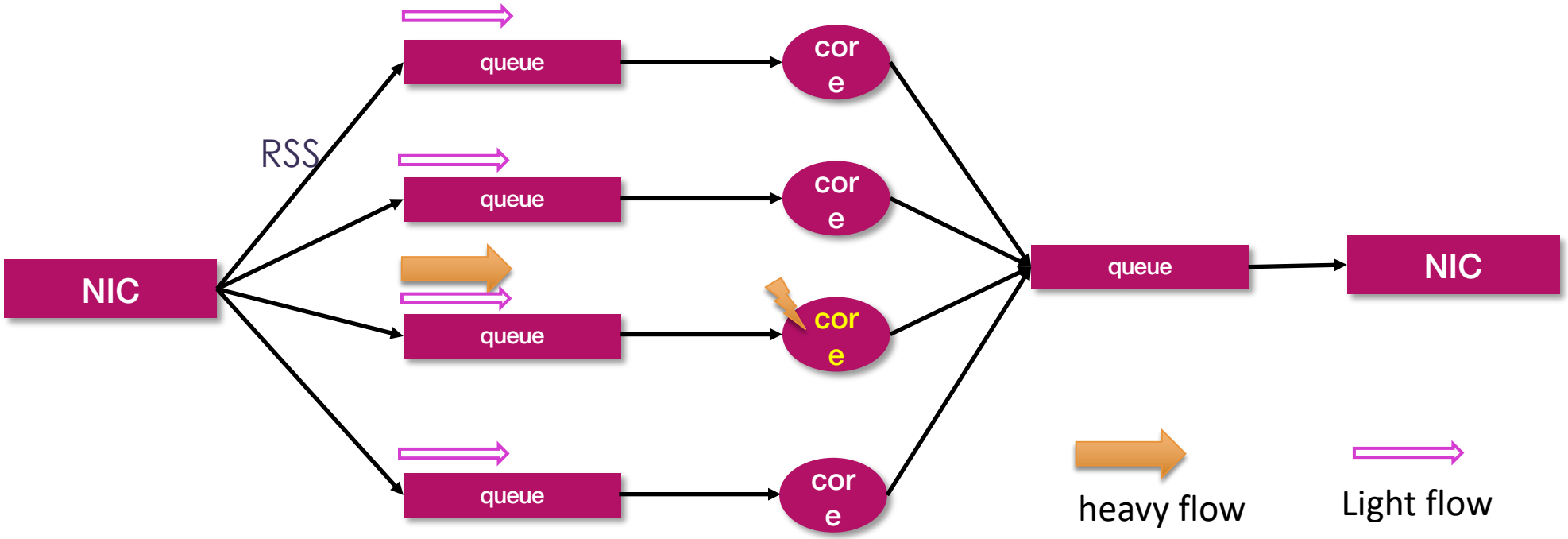
Handle 10Mpps Level Elephant Flow

Address the business challenge of Elephant Flow

Up to 3x higher performance

Scaling for DNAT and SNAT

Problem Statement



- The load balancer itself processes flows using multiple cores.
- NIC uses RSS to distribute flows among cores.
- However, flows are not equal, 10% elephant (heavy) flows may take 90% of total traffic.
- The small number of elephant flows may not be balanced by RSS.

Problem Statement (Cont.)

DPDK SUMMIT APAC•2021

- Need to treat the elephant flows differently than mice flows.
  - Huge volume of traffic from an elephant flow could exceed t single core's processing capacity.
  - Heavy flows and mice flows could impact each other from QoS perspective.

- What we propose to solve the issues.
  - Step 1: An Efficient heavy hitter detection Algorithm.
  - Step 2: Distributes a single elephant flow to multiple cores for parallel processing.
  - Step 3: Reorders the paralleled flow among multiple cores.

Key Components of Solution

DPDK SUMMIT APAC•2021

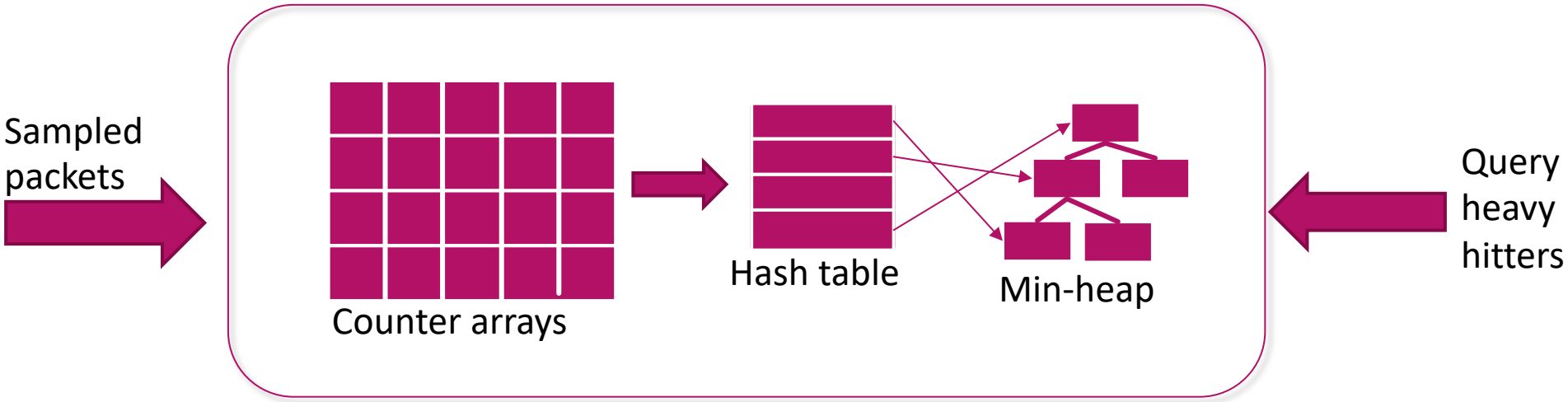
- Heavy hitter detection algorithm
  - Based on the state-of-the-art heavy flow detection algorithm - Nitrosketch [1].
  - Implemented and optimized for Intel Platform.

- Intel® Dynamic Load Balancer (DLB)
  - DLB is Intel's new hardware accelerator for queue management and load balancing.
  - We use DLB to distribute heavy flow among multiple cores.

- The processing pipeline of HDSLB

[1] Zaoxing Liu, et al. Nitrosketch: robust and general sketch-based monitoring in software switches (SIGCOMM '19)

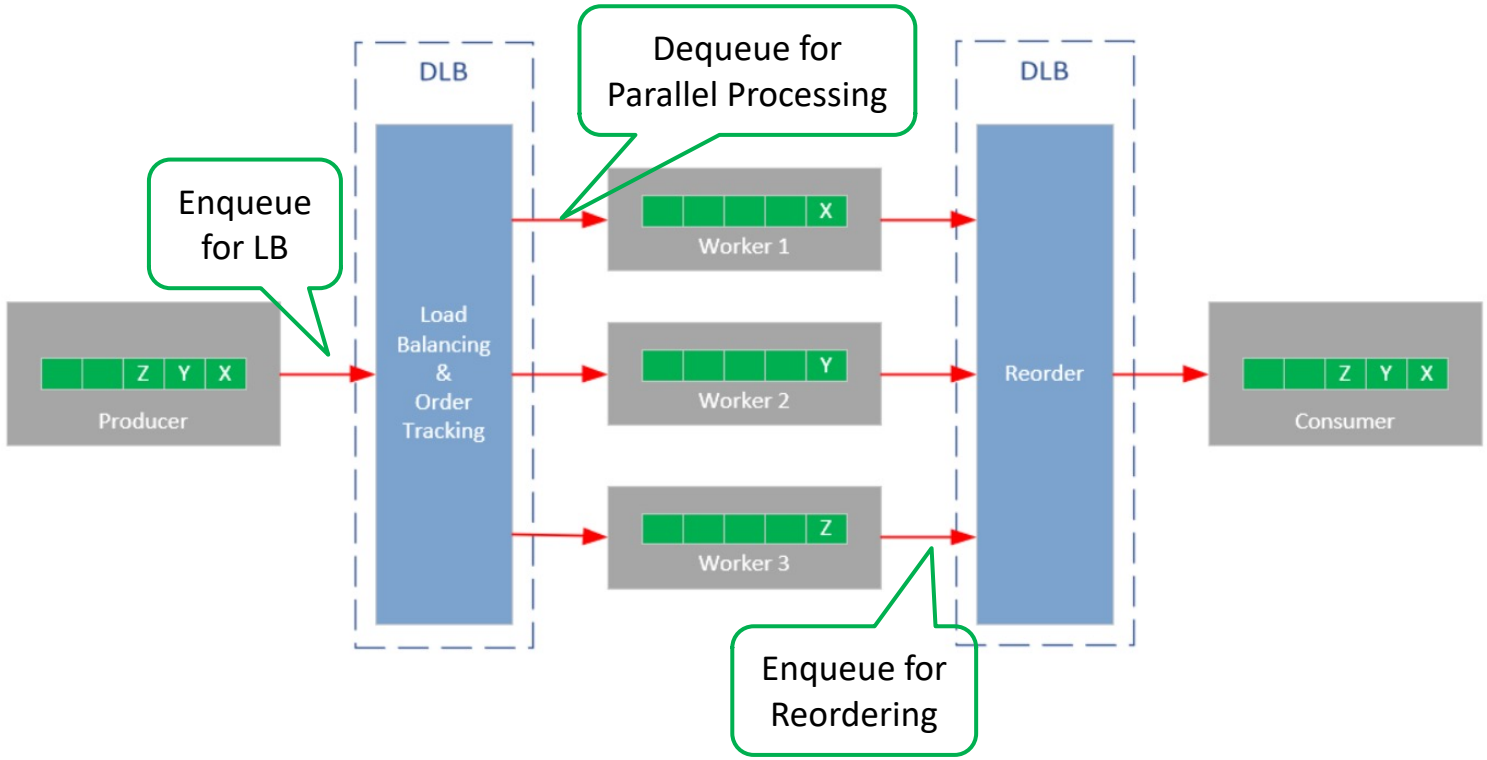
Heavy Hitter Detection Algorithm



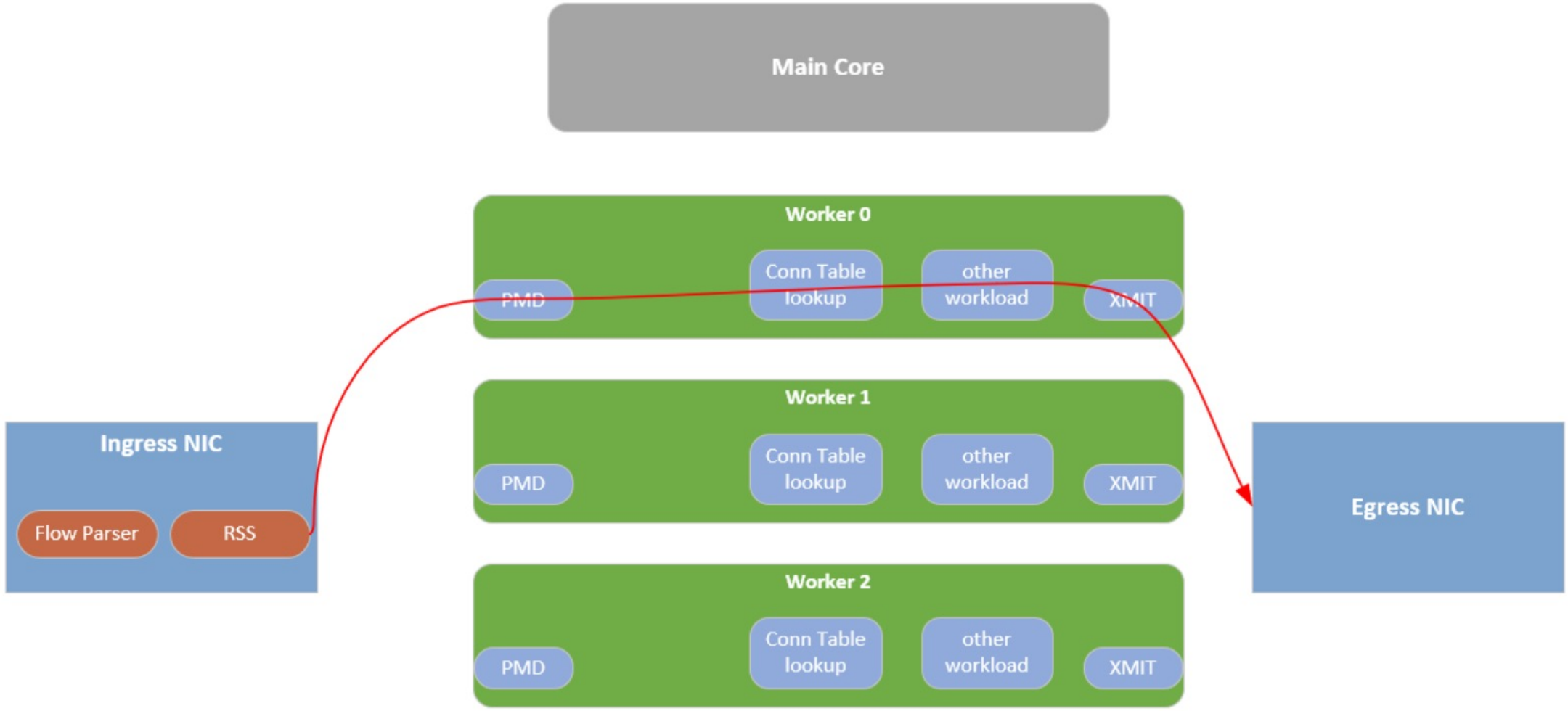
- The algorithm profiles and reports heavy flows with their estimated packet counts.
- The data structure is small enough to reside in local cache.
- Only a small percentage of total packets needs to be sampled (e.g. 1%, configurable).
- Uses a hash table to optimize the heap lookup time.
- Collaborating with Professor Liu, the author of Nitrosketch to further improve the algorithm.

Intel® DLB

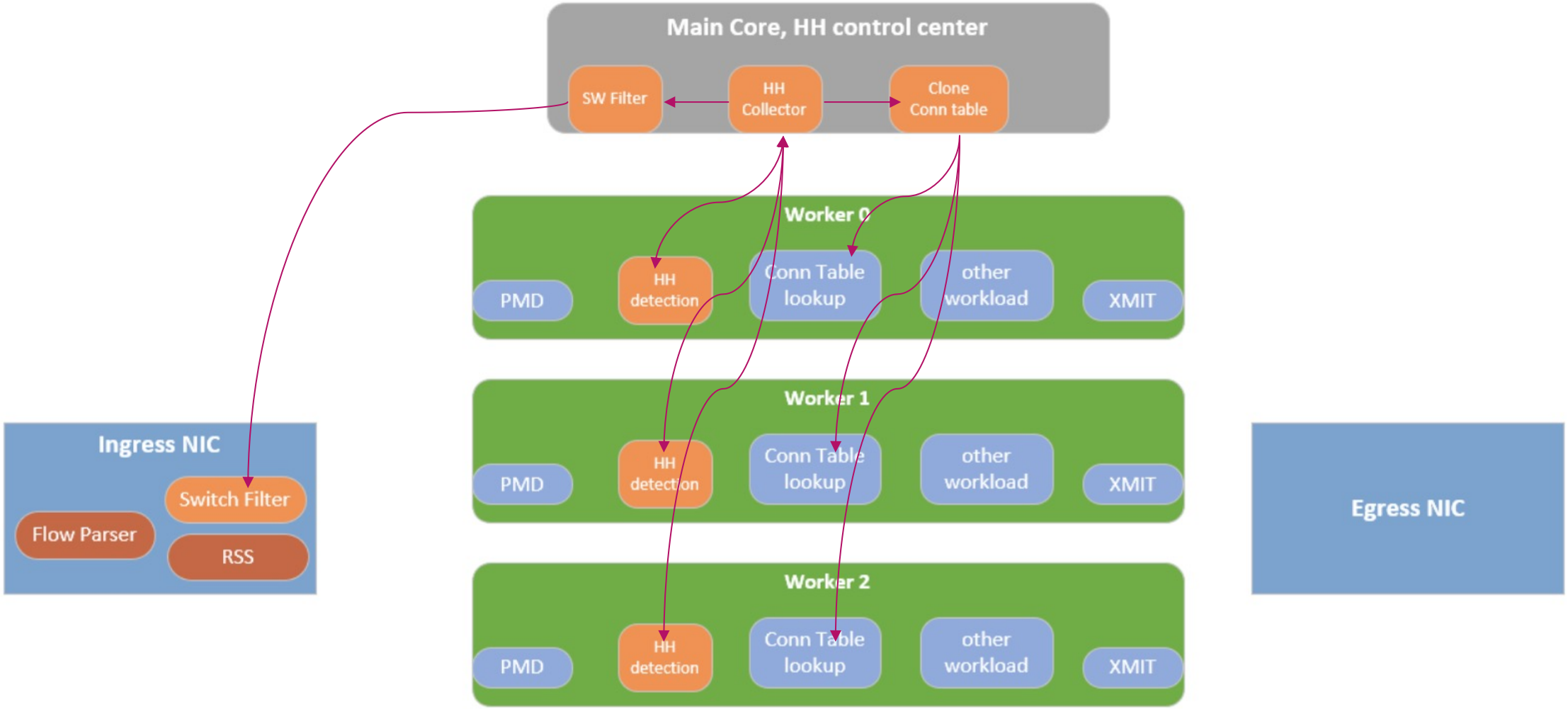
- Intel® DLB is a hardware accelerator available in Intel's latest processor
  - Dynamic Load Balancing
  - Exposes as a PCIe device
  - Acts as an event-dev in DPDK
  - A variety of working modes
- We use the Intel® DLB to distribute heavy flows among multiple cores.



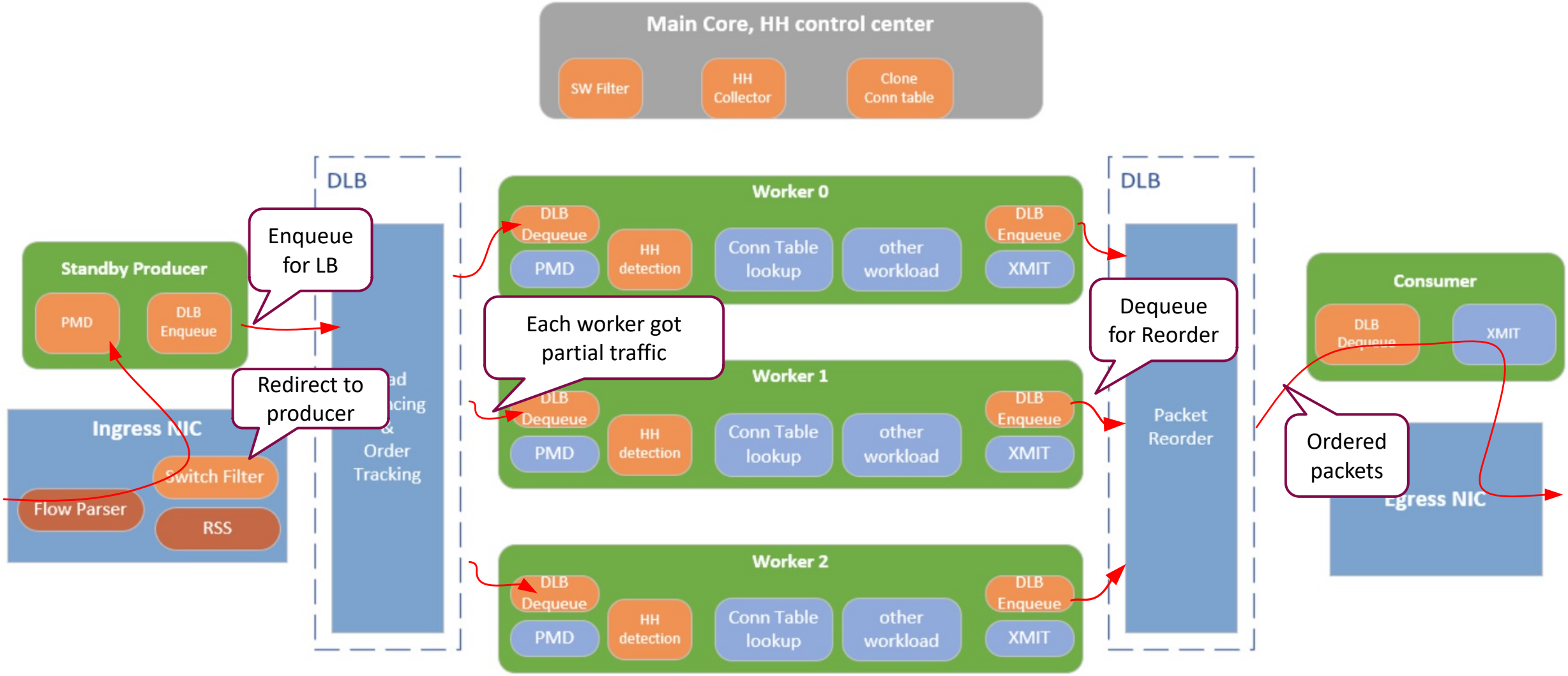
Mice Flows Processing



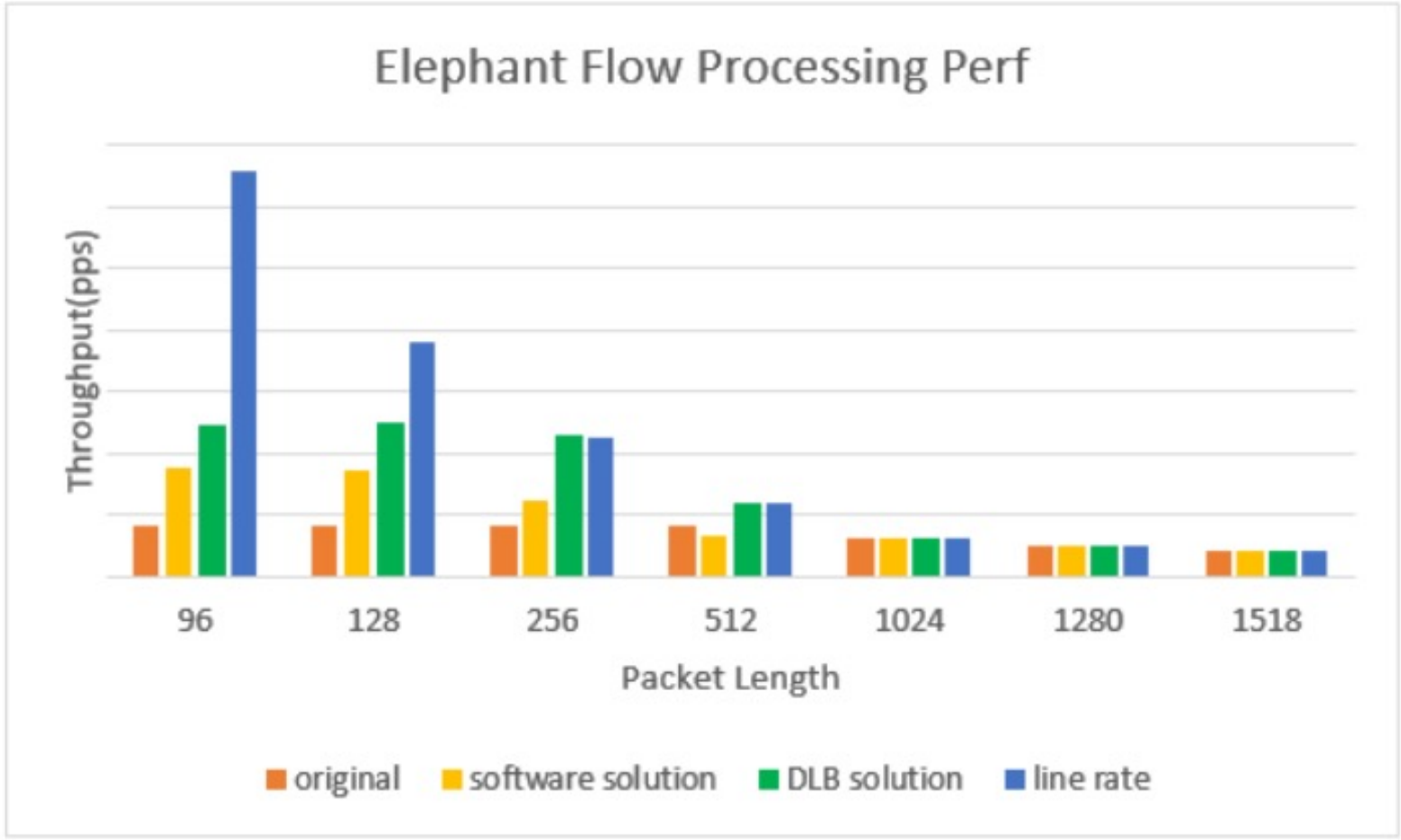
Elephant Flow Processing – Main Core



Elephant Flow Processing – Worker Cores



Performance Comparison



Key Takeaway

DPDK SUMMIT APAC•2021

- Distributing elephant flows to multiple cores is essential.

- Implements flow detection and distribution mechanism in HDSLB.
  - Improved state-of-the-art elephant flow detection algorithm
  - Leverages Intel® DLB technology to further reduce overhead.

- Other learnings
  - Leverages dedicated packet pools to avoid side effects on other workers.
  - Prefetch/CLDEMOTE instructions to hide the cache misses.
  - Leverages NIC offloading capability to accelerate packet processing.

Acknowledgement

- Jay Vincent @ Intel
- Pan Zhang @ Intel
- Mrittika Ganguli @ Intel
- Rahul R Shah @ Intel
- Niall McDonnell @ Intel
- Pravin Pathak @ Intel
- Sameh Gobriel @ Intel labs
- Ren Wang @ Intel Labs
- Charlie Tai @ Intel Labs
- Alan (Zaoxing) Liu @ Boston University

Thank You!

Q & A


DPDK
SUMMIT
APAC•2021