OVS-DPDK Performance Benchmark and Analysis with Multi-VMs

------- Last Level Cache ( LLC ) Part

Yao Lei (Lei.a.yao@intel.com)
For OVS-DPDK deployment with multi VMs, memory copy by vhost enqueue and dequeue cost large part of the CPU cycles. When packet size is large, LLC miss ratio is very important for memory copy efficiency.

Following two parts will impact the LLC hit ratio:

- DDIO: LLC miss caused by DDIO capacity
- Noise Neighbor: Workload on in VM
Intel® Data Direct I/O Technology (Intel DDIO) is a platform technology that enables I/O data transfers that require far fewer trips to memory. RX side as following:
1. **DDIO: Calculation the Cache Budget**

Take Intel Xeon(R) Platinum 8180 as example. LLC size is 38.5MB, when packet size=1518B, the total packet can be hold in DDIO capacity ≈ \( \frac{38.5MB \times 10\%}{1518} = 2656 \)

Default rxd number of OVS-DPDK is 2048, is it suitable for all scenario and settings?
2. Noise Neighbor

All the cores on the same socket will share the LLC. Workload such as memory r/w in VM will impact the LLC miss ratio of OVS core

Settings:
OVS-DPDK (4 core). 32 local VM(16C/32T)
Noise-neighbor: stress-ng --vm-rw

CPU: Intel(R) Xeon(R) CPU E5-2699 v4 @ 2.20GHz
LLC Size: 55MB
LLC Ways: 20
Intel Cache Allocation Technology (CAT) helps address shared resource concerns by providing software control of where data is allocated into the last-level cache (LLC).
How much LLC way should be dedicated to OVS-DPDK?

Sample command for CAT:
./pqos -e "llc:0=0x000ff"
./pqos -e "llc:1=0xffff00"
./pqos -a "llc:0=1-4"
./pqos -a "llc:1=0,5-21"

Dedicate the first 8 LLC ways to OVS core(1-4)
Conclusion:

LLC Miss Ratio will impact the OVS-DPDK throughput when packet size is large. Two optional way to avoid this situation:

1. Keep the total NIC RX descriptor number within the Processor's DDIO budget.
2. Dedicate part of the LLC to the OVS-DPDK core to avoid noise neighbor impact

Thanks!