

DPDK nad sieťovými kartami COMBO

Matej Vido

<xvidom00@stud.fit.vutbr.cz>

Fakulta informačních technologií, Vysoké učení technické v Brně

V spolupráci s



DPDK princípy

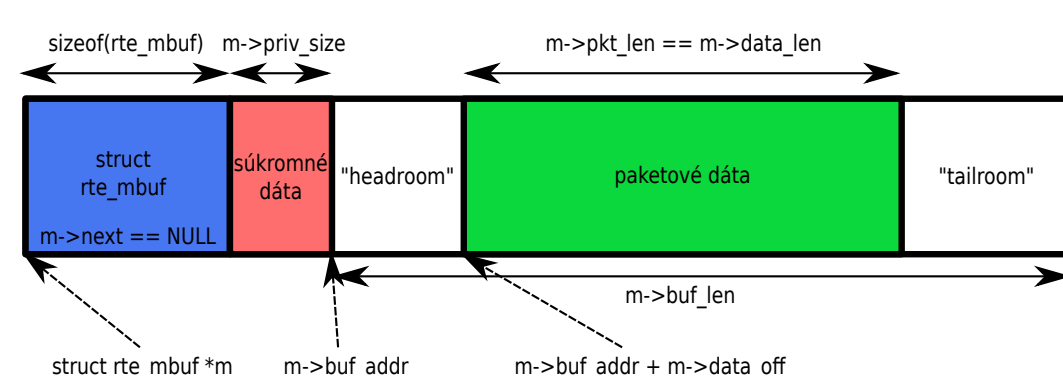
Ovládače v DPDK dosahujú zvyšovanie výkonnosti nasledujúcimi technikami:

- > dopytovací režim (poll mode)
- > príjem a odosielanie zhukov paketov
- > dopredu alokované buffere
- > obrovské stránky (huge pages)
- > bezzámkové spracovanie
- > rozdelenie spracovania paketov medzi jadrá CPU
- > nulové kopírovanie dát

Pakety sú ukladané v bufferoch nazývaných *mbuf*. Úlohou ovládača je naplniť pole mbufov prijatými paketmi a odosielať pakety z mbufov do siete.

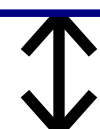
DPDK mbuf

Mbufy sú spravované správcom pamäťových objektov a majú pevnú veľkosť. Pakety, ktoré sú väčšie ako voľné miesto v mbufe, sa ukladajú do zretazených mbufov.



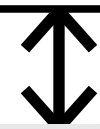
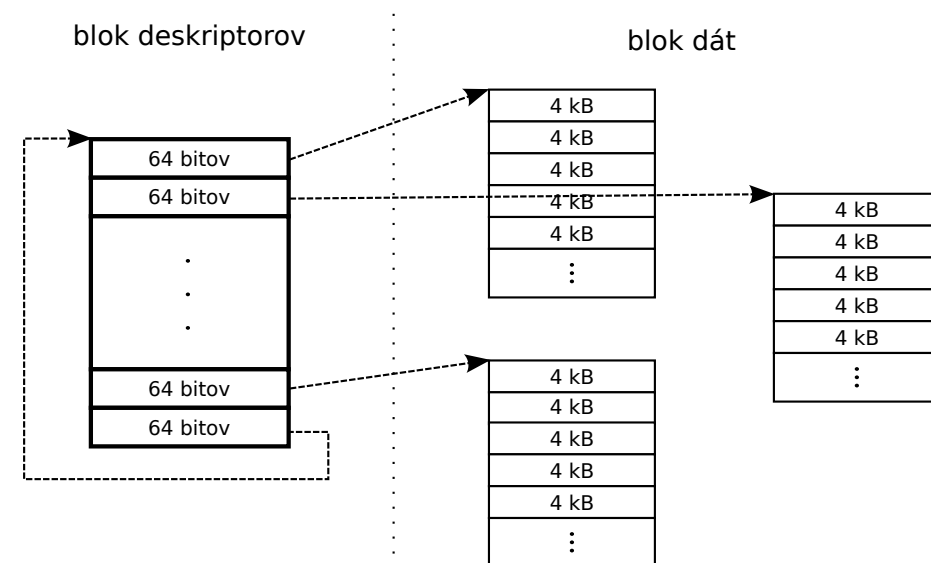
DPDK ovládač szedata2

Pakety v bufferi používanom rozhraním SZE2 sa náchádzajú tesne vedľa seba. Prenos paketov v súvislých pamäťových blokoch znižuje réžiu prenosu dát po zbernici. Rozhranie SZE2 je vhodné pre prúdové spracovanie dát. Ovládač szedata2 kopíruje dáta medzi mbufmi a SZE2 kruhovými buffermi, vďaka čomu je možné nad ovládačom postaviť DPDK aplikácie bez obmedzení kladených rozhraním SZE2 a zároveň využiť prenosové pásmo zbernice efektívne aj pre krátke pakety.



SZE2 rozhranie

Zabezpečuje DMA prenosy dát po zbernici PCI-Express 3.0 medzi COMBO kartou a RAM v počítači. Rozhranie je tvorené modulmi jadra a knižnicou **libsze2**. Dáta sú prenášané do kruhového bufferu napájaného do užívateľského priestoru tvoreného niekoľkými veľkými pamäťovými blokmi, v ktorých sa pakety náchádzajú tesne vedľa seba.

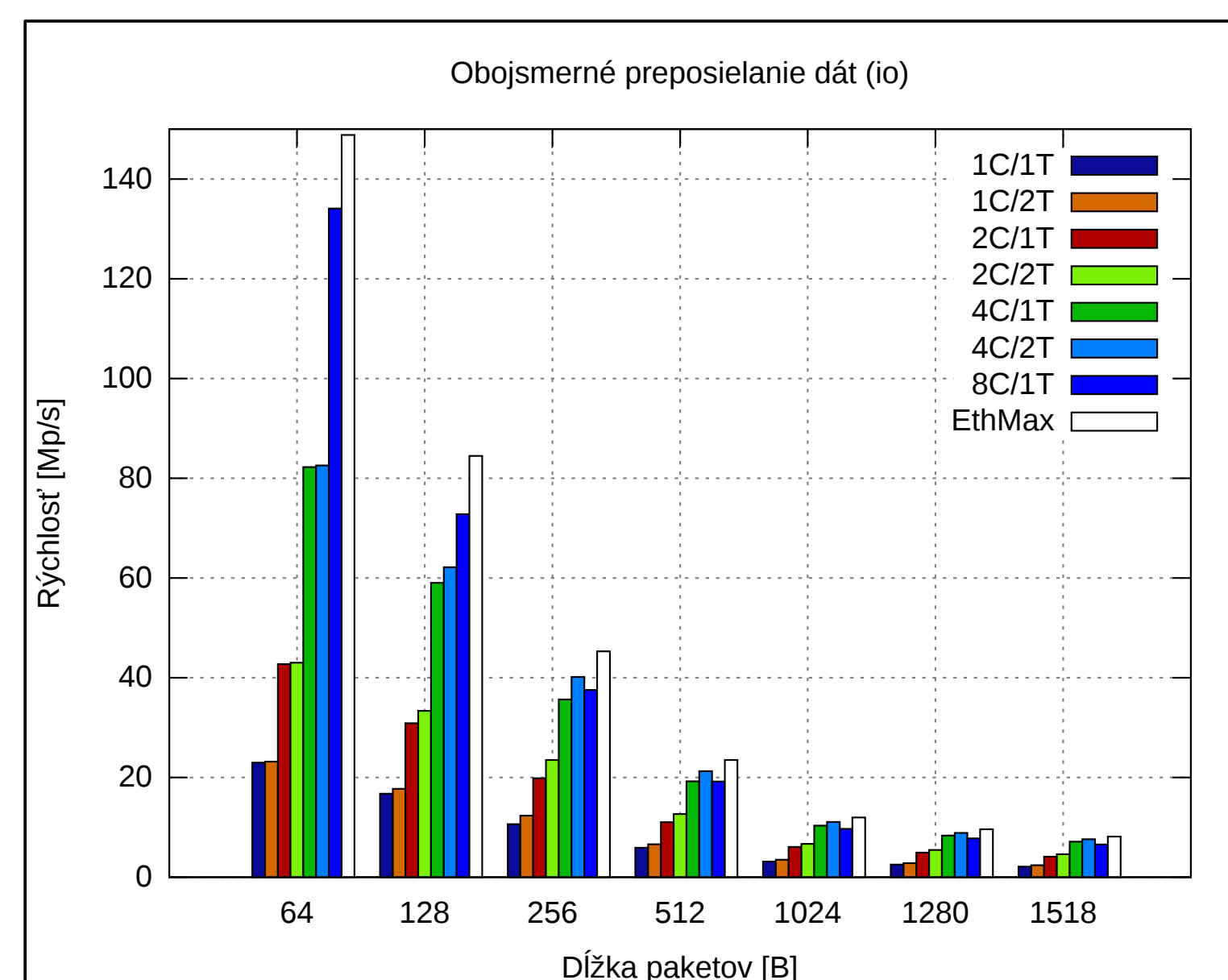
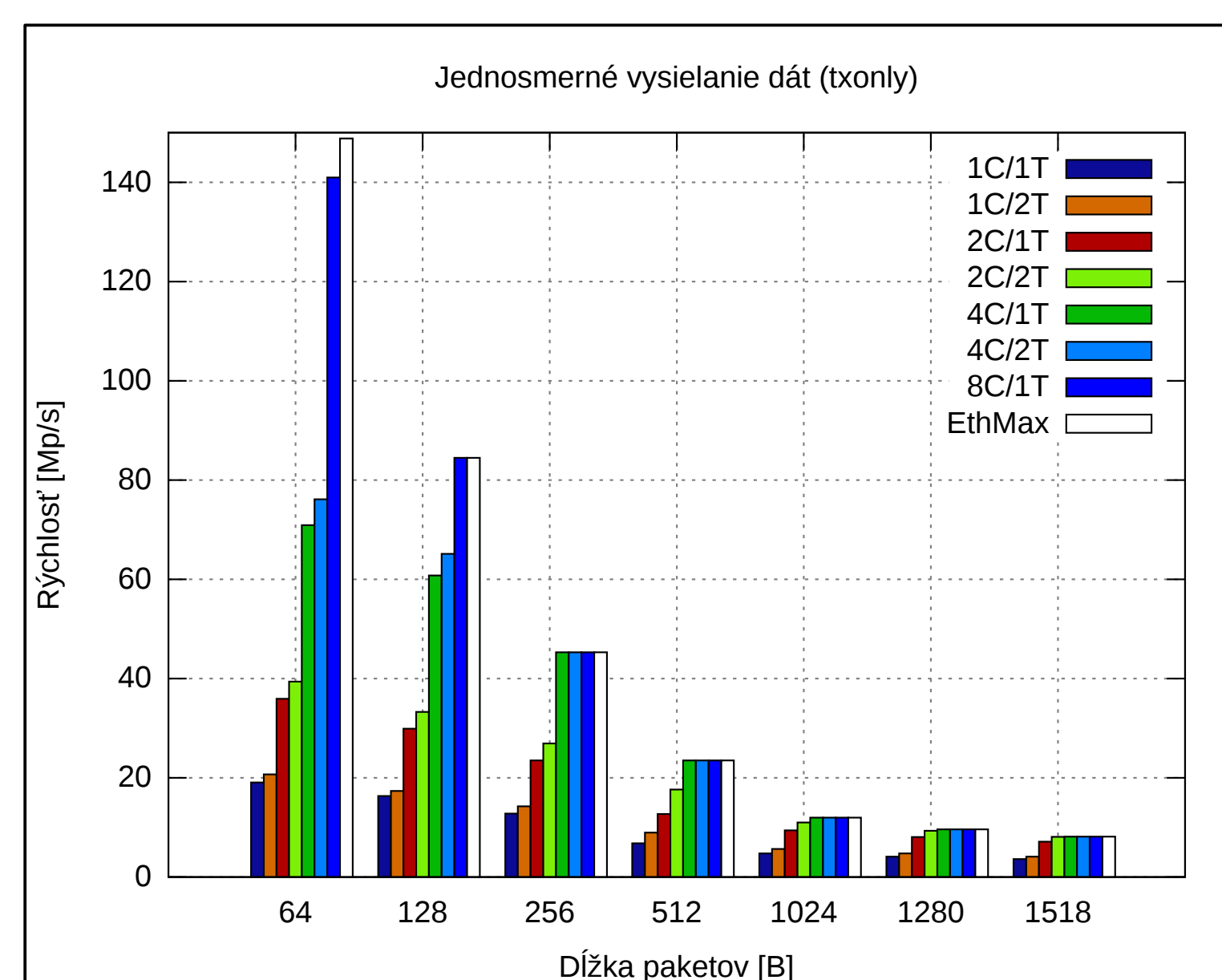
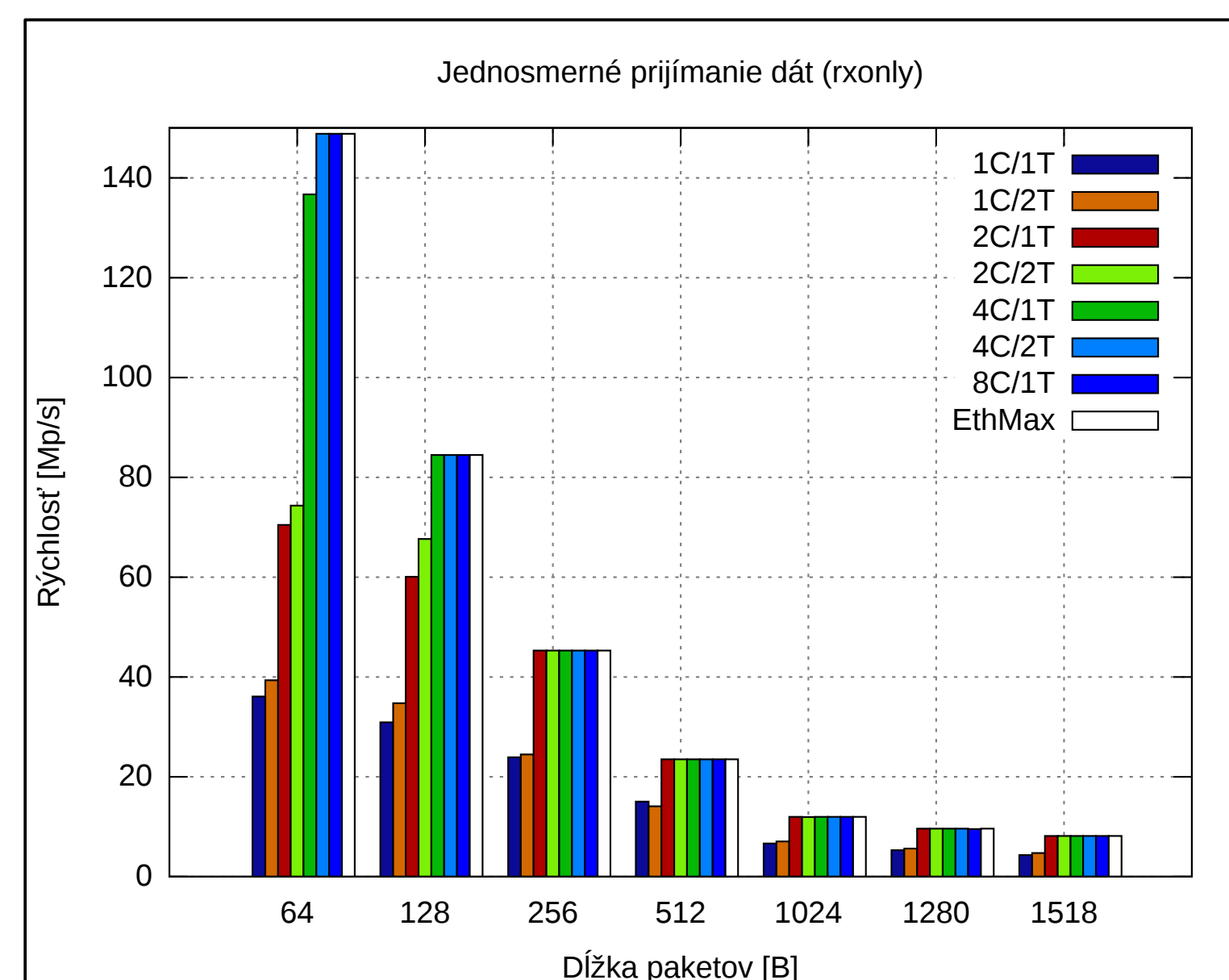


COMBO-100G¹



Výsledky meraní

- 1C/1T - 1x DMA kanál 1 fyzické jadro
- 1C/2T - 2x DMA kanál 1 fyzické jadro (+ hyperthreading)
- 2C/1T - 2x DMA kanál 2 fyzické jadrá
- 2C/2T - 4x DMA kanál 2 fyzické jadrá (+ hyperthreading)
- 4C/1T - 4x DMA kanál 4 fyzické jadrá
- 4C/2T - 8x DMA kanál 4 fyzické jadrá (+ hyperthreading)
- 8C/1T - 8x DMA kanál 8 fyzických jadier



¹ zdroj: <https://www.liberouter.org/cards-gallery/>