

PACKET CAPTURE & REPLAY ON 100 GB/S ETHERNET

Václav Hummel

Bachelor Degree Programme (3), FIT BUT

E-mail: xhumme00@stud.fit.vutbr.cz

Supervised by: Jiří Matoušek

E-mail: imatousek@fit.vutbr.cz

Abstract: Future backbone networks will operate at speed of hundreds of Gb/s. Network devices such as routers or switches have to be tested properly at full wire speed before deployment. This implies the need of packet capturing and replaying at this speed. Software solutions are inconvenient. Therefore, this paper proposes hardware accelerated application for packet capturing and replaying, which is based on the NetCOPE platform and the COMBO-100G card. This solution guarantees throughput of 100 Gb/s per each network interface. Packet emission can be driven either by precise 64-bit timestamps for time critical experiments or by preset traffic speed.

Keywords: 100 Gigabit Ethernet, packet capture & replay, precise timestamps, NetCOPE

1 ÚVOD

Fenomén Internetu již zasáhl celou společnost. S tím jak roste jeho popularita, rostou i nároky na něj kladené. Je to dáno tím, že této technologii jsou svěřovány věci dříve nevídané - peníze a dokonce i lidské životy. Jedním z parametrů na páteřních linkách v rámci Internetu je přenosová rychlost. V dnešní době je zde běžně nasazena technologie 10 Gb/s Ethernetu. V blízké budoucnosti se počítá s řádově větší přenosovou rychlostí.

Ještě dříve než se síťová infrastruktura a její jednotlivé prvky uvedou do ostrého provozu, je nutno je řádně otestovat na plné rychlosti linky. Testování spočívá v přehrávání nebo generování síťového provozu, který se posílá na vstup testovaného zařízení. Výstup tohoto zařízení se zachytává pro další analýzu. Tento problém lze řešit třemi způsoby: softwarovými nástoji, speciálním hardwarovým testovacím zařízením nebo pomocí přídatné karty s programovatelným FPGA čipem do běžných serverových stanic.

První přístup lze realizovat pomocí volně dostupných softwarových nástrojů (`tcpdump` pro zachytávání a `tcpreplay` pro přehrávání) a vyniká tak svou dostupností. Bohužel však toto řešení není schopno dosáhnout plné rychlosti linky a nelze zaručit přesný čas odeslání dat [4]. Specializované hardwarové zařízení oba problémy řeší, ale výrobce si je nechá také patřičně zaplatit, což je činí často nedostupnými. Alternativou je řešení založené na technologii FPGA. Časově kritické operace jsou zde implementovány přímo v hardware, a je tak zajištěná dostatečná rychlost zpracování. Přitom cena přídatných FPGA karet je oproti specializovanému hardware řádově nižší.

Na posledním uvedeném přístupu je založen systém pro zachytávání a přehrávání paketů, jehož architekturu popisuje tato práce. Toto řešení je navrženo pro kartu COMBO-100G a platformu NetCOPE, které jsou představeny v kapitole 2. V kapitole 3 je pak popsána samotná architektura systému pro zachytávání a přehrávání paketů. Vysílání může být řízeno pomocí přesných časových značek (časově kritické aplikace), nebo pomocí předem přednastavené rychlosti linky.

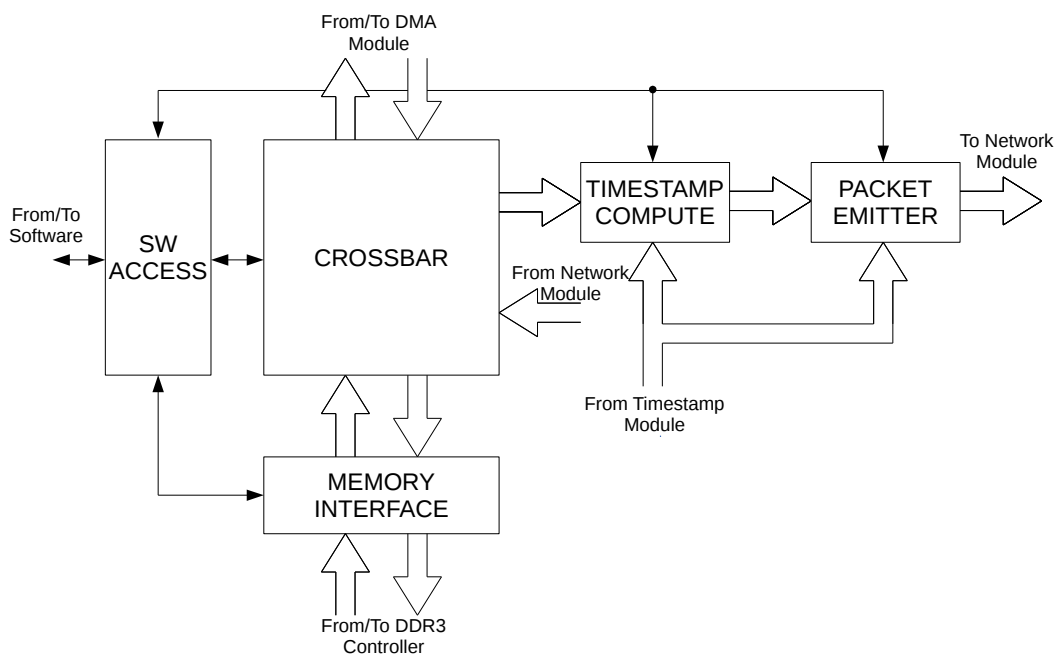
2 KARTA COMBO-100G A PLATFORMA NETCOPE

Jádrum akcelerační karty COMBO-100G je FPGA čip Virtex-7 HT od firmy Xilinx. Karta obsahuje jeden optický modul CFP2, který je součástí fyzické vrstvy 100 Gb/s Ethernetu. Dostatečně široké přenosové pásmo mezi akcelerační kartou a hostitelskou stanicí zajišťují dvě koncová zařízení (*end-points*) PCIe x8 generace 3 na čipu FPGA. Na kartě jsou osazeny 3 statické paměti typu QDRIIIe a 8 dynamických pamětí typu DDR3 [1]. Dále je zde hardwarová podpora pro generování signálu PPS (*pulse per second*), který je důležitý pro generování přesných časových značek [3].

Platforma NetCOPE slouží pro rychlý vývoj hardwarově akcelerovaných síťových aplikací na kartách rodiny COMBO [2]. Její součástí je softwarová a firmwarová vrstva. Softwarová vrstva obsahuje ovladače, knihovny a podpůrné nástroje, které mj. umožňují pracovat se soubory typu PCAP. Firmwarová vrstva implementovaná na čipu FPGA se skládá ze síťového modulu, modulu připojovacího sběrnici PCIe, DMA modulu, paměťových řadičů a modulu pro generování časových značek s přesností na nanosekundy.

3 ARCHITEKTURA SYSTÉMU PRO ZACHYTÁVÁNÍ A PŘEHRÁVÁNÍ PAKETŮ

Navržená architektura systému pro zachytávání a přehrávání paketů pro jedno síťové rozhraní je zobrazena na obrázku 1. Toto řešení je škálovatelné na více rozhraní pouhým zopakováním schématu.



Obrázek 1: Architektura systému pro zachytávání a přehrávání paketů pro jedno síťové rozhraní

Architektura se skládá z jednotek SW ACCESS, MEMORY INTERFACE, TIMESTAMP COMPUTE, PACKET EMITTER a CROSSBAR. Jednotka SW ACCESS slouží pro obousměrnou komunikaci s hostitelskou stanicí. Je to aplikačně specifický procesor s velmi jednoduchou instrukční sadou a stavovými a konfiguračními registry, který řídí celý systém na základě požadavků ze softwaru. Jednotka MEMORY INTERFACE tvoří rozhraní mezi navrženým systémem a externí pamětí na kartě. Pro první funkční prototyp již bylo navrženo, implementováno a verifikováno rozhraní pro statickou paměť QDRIIIe. Výsledky syntézy shrnuje tabulka 1. Velmi důležitá je dosažená frekvence přes 300 MHz (požadováno je alespoň 200 MHz). U finální verze se počítá s použitím řádově větší dynamické paměti, kdy postačí vyměnit pouze tuto jednotku, nikoliv celý systém. Jednotka TIMESTAMP COMPUTE vypočte před odesláním paketů nové časové značky pro řízení odeslání tak,

že ke všem časovým značkám v zachyceném provozu připočte stejnou konstantu odpovídající rozdílu mezi první časovou značkou v zachyceném provozu a aktuálním časem. Původní časové rozestupy mezi zachycenými pakety tak zůstanou zachovány. Jednotka PACKET EMITTER posílá pakety do síťového modulu s respektováním časových značek, nebo přednastavené rychlosti odesílání, a nebo zcela bez omezení. Poslední jednotkou je CROSSBAR, který dynamicky propojuje datové cesty.

Čip	LUT	Registry	BRAM	Frekvence [MHz]	Nástroj
XC7VH580T	5736 (1 %)	3643 (0 %)	26 (2 %)	300,728	XST 14.7

Tabulka 1: Výsledky syntézy rozhraní pro statickou paměť QDRIIIe, v závorkách využití čipu

Výše popsaný systém podporuje režimy zachytávání nebo přehrávání síťového provozu a režim běžné síťové karty, mezi kterými se může dynamicky přepínat. Zachytávání je podporováno přímo ze sítě prostřednictvím síťového modulu, a nebo ze softwaru ze souboru typu PCAP. Zachycený síťový provoz je uložen do dynamické paměti na kartě. Z této dynamické paměti je možno provoz přehrát na síťové rozhraní, nebo do softwaru do souboru typu PCAP. Vysílání na síťové rozhraní řídí jednotka PACKET EMITTER, a je tak možno vysílat pakety s respektováním časových značek, nebo přednastavené rychlosti odesílání, a nebo zcela bez omezení. Záleží na aktuálním stavu jednotky SW ACCESS, která je konfigurovatelná ze softwaru. Režim běžné síťové karty spočívá ve správném propojení datových cest jednotkou CROSSBAR. Z vlastností systému vyplývá ještě čtvrtý režim paketového generátoru. Syntetický provoz se vytvoří v softwaru ve formátu PCAP, pošle se do dynamické paměti na kartě a je následně přehrán na výstupní síťové rozhraní.

4 ZÁVĚR

Tento článek představuje architekturu systému pro zachytávání a přehrávání paketů na 100 Gb/s Ethernetu, který je založen na platformě NetCOPE a kartě COMBO-100G. Pro zachytávání i přehrávání se používá dynamická paměť typu DDR3 přímo na kartě. Vysílání paketů je možno řídit pomocí přesných časových značek, nebo předem přednastavené rychlosti linky. Díky podpoře souborů typu PCAP je možno řešení použít i jako paketový generátor. Řešení garantuje přesný čas odeslání dat a to vždy na plné rychlosti linky.

PODĚKOVÁNÍ

Článek vznikl za podpory projektu VUT v Brně FIT-S-14-2297 a projektu TA ČR TA03010561.

REFERENCE

- [1] Friedl, Š., Puš, V., Matoušek, J., Špinler, M. *Designing a Card for 100 Gb/s Network Monitoring*. Technická zpráva 7. CESNET z. s. p. o., 2013.
- [2] Martínek, T., Košek, M. *NetCOPE: Platform for Rapid Development of Network Applications*. In 11th IEEE Workshop on Design and Diagnostics of Electronic Circuits and Systems. DDECS 2008, 2008. s. 1-6.
- [3] Martínek, T., Žádník, M. *Precise Timestamp Generation Module and its Applications in Flow Monitoring*. Technická zpráva 13. CESNET z. s. p. o., 2009.
- [4] Covington, G. A., Gibb, G., Lockwood, J., aj. *A Packet Generator on the NetFPGA Platform*. In 17th IEEE Symposium on Field Programmable Custom Computing Machines. FCCM '09, 2009. s. 235-238.