

PDS – Testování propustnosti síťového stacku,

Pavel Koupý - xkoupý00

Zpřesnění zadání

Projekt je tedy mířen na rozdíl mezi implementací směrování paketů pomocí linuxového síťového stacku a implementací směrování pomocí technologií XDP technologie a důsledků použití na mikropočítači Raspberry Pi 3. XDP je určena k akceleraci síťových směrovačů k tomuto účelu vyrobených a se síťovými rozhraními o rychlostech desítek Gbps a více, ale jakého zlepšení schopno dosáhnout u ARMového mikropočítače se čtyřmi jádry a ethernetovým čipem připojeným přes USB rozhraní v rámci interního i externího rozhraní Ethernet.

Technologie XDP

Použití eXpress Data Path představuje možnost zpracovat paket na velice nízké úrovni, ne-li přímo v ovladači síťového zařízení a vyhnout se tak alokaci sk_buff struktury, které je nezbytná pro další práci s paketem v systému.

Technologie XDP [2,3] je svázána s Berkeley Packet Filter (konkrétně extended BPF) tím způsobem, že poskytuje možnost dynamicky načíst a spustit verifikovaný XDP bytekód, za použití in-kernel Just-In-Time kompilátoru, který BPF bytekód přeloží do nativní opkódu, hluboko v systémovém jádře nad síťovými hooks, konkrétně zde XDP hook.

Při zpracování paketů pomocí XDP je paket ohodnocen akcí, která určí jeho další průchod systémem. XDP umožňuje paket postoupit systému a pokračovat ve standardním postupu síťovým stackem pomocí akce XDP_PASS, nebo co je zajímavější umožňuje paket přesměrovat pomocí XDP_REDIRECT. Další možnost je poslání zpět stejným síťovým rozhraním pomocí XDP_TX, vhodné například pro vyvažování zátěže. XDP_DROP paket zahodí a používá se zejména pro implementaci firewallu či mitigace DDoS útoku.

Metodologie

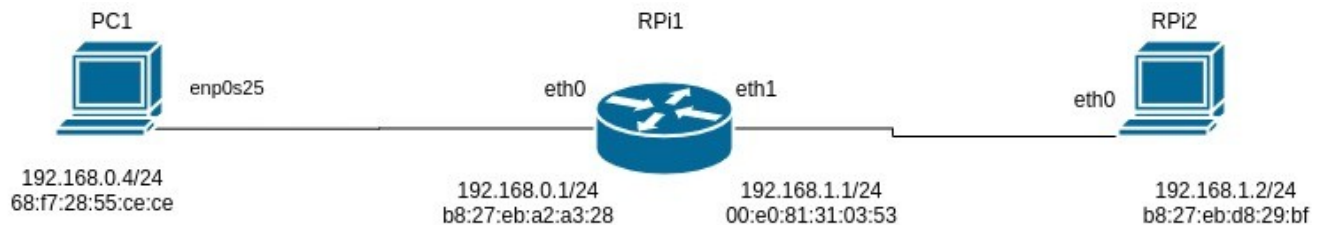
Propustnost síťového uzlu je možné měřit dle počtu zpracovaných bitů za sekundu (bits per second – bps) nebo jako počet zpracovaných paketů za sekundu (packet per second - pps). Počet paketů obvykle lépe modeluje možnosti daného systému v rámci jednotek zpracovávajících pakety, tedy například v tomto projektu se testování soustředí na mikropočítač Raspberry Pi, v tom smyslu, že počet paketů za sekundu svědčí o konkrétním zpracování jádrem systému, kdežto počet zpracovaných bitů za sekundu, která je dnes udávána v stovkách Mega až Giga, spíše ukazuje na rychlost se kterou pracují samotné síťové karty.

Pro měření je použit jednoduchý kód z článku „How to receive a milion packets“, který se sestává ze dvou částí. Sender posílá UDP pakety ve zvoleném počtu se zvoleným paylodem na danou adresu a

port. Druhou částí je přijímač těchto paketů a výpočet množství zpracovaných paketů je výstupem přijímače a tento výstup je použit pro vytvoření grafu.

Topologie

Topologie je jednoduchá a je postavena ze tří linuxových uzlů, kdežto Raspberry Pi figuruje jako směrovač RPi1 a zachytávací zařízení RPi2, osobní počítač PC1 jako generátor dat.



Obrázek 1: Použitá topologie

Měření bylo provedeno ve třech scénářích.

1. scénář je následující, v prvním kroku je RPi2 nastavena statická cesta do sítě 192.168.0.0/24 a poté spuštěn program udpreceiver, toto zařízení bude sloužit k zachytávání paketů. RPi1 bude sloužit jako směrovač a pro směrování v rámci linuxového jádra je zapotřebí zapnout ip_forwarding tedy směrování paketů. PC1 bude sloužit jako generátor udp paketů a i zde je staticky přidána cesta do sítě přijímače RPi2, 192.168.1.0/24.
2. scénář je měření s použitím XDP a volání bpf_redirect(). Topologicky je situace obdobná s prvním scénářem, ale tímto případem zakážeme ip_forwarding v systému a použijeme XDP hooks k odchyzení všech paketů a statickému přesměrování na daný interface s použitím konstantní cílové MAC adresy.
3. scénář slouží především k porovnání rychlostí zpracování v rámci linuxového síťového jádra. Ukazuje jakou rychlostí je schopno Raspberry Pi zpracovávat data bez směrování tedy přes přímý spoj. PC1 slouží opět jako generátor a RPi1 jako přijímač dat.

Výhodou použití UDP komunikace v tomto případě je, že není nutné se starat o zpětné směrování odpovědí a potvrzení od protějšky strany a nevytváří se stavová komunikace. Vzhledem k tomu je možné použít jednoduchý příklad z tutorialu k XDP [3] pro statické směrování. Jen měříme počet zpracovaných paketů zaslaných jedním procesem druhému procesu, kde důležitou proměnnou bude tvořit směrovací technologie na mikropočítači Raspberry Pi, který je použit při měření jako směrovač. Pakety obsahují payload 32 bajtů a celkově má paket na Ethernetu 74 bajtů a je jich zasláno 1024 najednou.

HW a prostředí

Je použito mikropočítače Raspberry Pi 3 Model B V1.2 (RPi1 a RPi2), který disponuje 4jádrovým ARMovým procesorem verze 8 Cortex-A53 a taktu 1,2Ghz. Operační paměť je 1 Gb a 10/100 Mbps

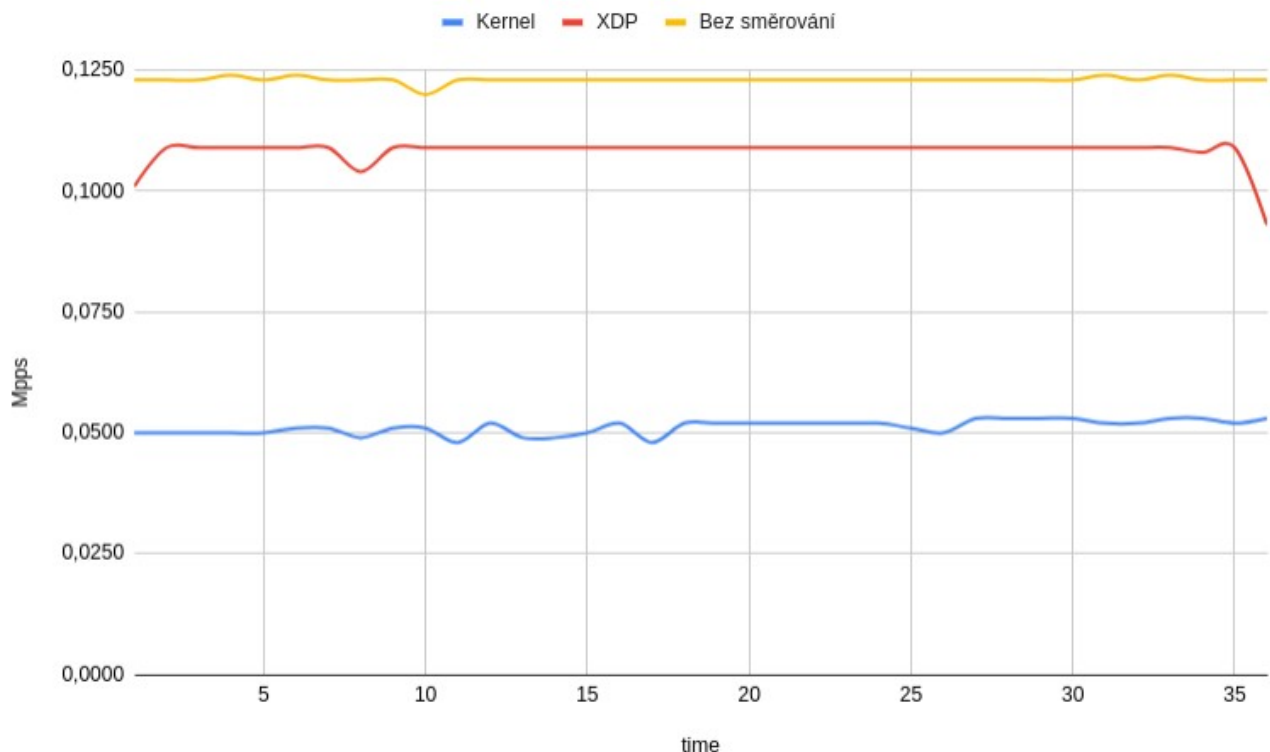
Ethernetem, který je napojen přes USB rozhraní. Osobní počítač PC1 je vybaven 4jádrovým Intel Core i5-5200U o taktu 2.2Ghz a 8 Gb operační paměti. Se síťovým rozhraním obdobně jako u Rpi 10/100 Mbps Ethernet, ale připojeného přes PCI sběrnici.

Pro použití XDP technologie na Linuxové distribuci Raspbian Buster, který je běžně dostupný, od vývojářů ve formě bitové kopie, pro mikropočítač Raspberry Pi . Nicméně je nutné přeložit kernel a přidat podporu pro načtení a překlad BPF bytekódu.

Kernel (v4.19.114-v7) lze přeložit na samotném Raspberry Pi , nebo pomocí cross-compileru na vlastním počítači. V síťových možnostech je nutné přidat možnost použití XDP sockets, ačkoli to není přímo nutné pro funkčnost tohoto měření. Dále podporu BPF frameworku, systémové volání bpf() a BPF Just-In-Time kompilér. Poté stačí znovu přeložit kernel a vytvořit potřebné součásti jako zImage, tedy obraz kernelu, moduly a Device Tree Blobs pro popis HW. Podrobnější postupy jsou v příložené dokumentaci.

Měření

V grafu níže je možné vidět srovnání technologie XDP oproti klasickému forwardingu či směrování pomocí síťového stacku Linuxu.



Obrázek 2: Graf rychlostí zpracování paketů

Naměřené hodnoty je možné porovnat s teoretickým maximem. Tedy pomocí nástroje iperf jsem změřil reálnou propustnost linky, která je 95Mbps a byly posílány udp pakety o velikosti 74 bajtů. Tedy teoretické maximum linky je pro takové parametry zhruba 0,158Mpps tedy je možné porovnat rychlost zpracování naměřenou při přímém spojení, která by této hodnotě měla být nejbližší a ta se skutečně někde pohybuje kolem 0,125Mpps. A rychlost XDP modulu je velice podobná.

Závěr

V projektu je ke zlepšení propustnosti využito XDP modulu, který je zařazen mezi síťové rozhraní a zpracování paketu v systému. Tento modul je dle měření rychlejší a propouští pakety skrze Raspberry Pi rychleji než při použití Linuxového síťového jádra. Rychlost XDP vykazovala rychlost podobnou přímému spojení a rychlosti bez směrování tedy maximální rychlosti bez další konfigurace systému, která je možná a pravděpodobně by vedla ještě ke zvýšení zpracovaných paketů. Nicméně použitý XDP modul není pro směrování vhodný pro správné směrování je třeba vytvořit odpovídající mapy v bpf uložisti a jeho naplnění s pomocí první pomalého průchodu systémem nebo statickými cestami uživatelem.

Reference

- [1] <https://blog.cloudflare.com/how-to-receive-a-million-packets/>
- [2] <https://www.iovisor.org/technology/xdp>
- [3] <https://github.com/xdp-project/xdp-tutorial>
- [4] <https://suchakra.wordpress.com/2017/05/23/an-entertaining-ebpf-xdp-adventure/>
- [5] <https://wis.fit.vutbr.cz/FIT/st/cfs.php.cs?file=%2Fcourse%2FPDS-IT%2Flectures%2Fpds-zpracovani-paketu.pdf&cid=13440>