Jak zajistit bezpečnost v datových centrech

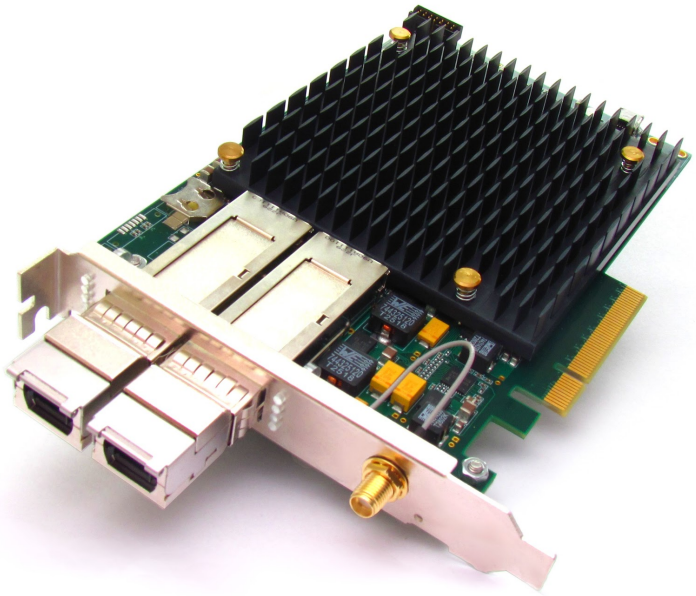
Málokdo ví, že v České republice existuje tým vědců, který vytváří specializovaný síťový hardware na světové úrovni založený na technologii FPGA již od roku 2003 [1]. Hlavní oblastí vědy a výzkumu je vývoj komplexních monitorovacích nástrojů pro vysokorychlostní sítě a datová centra. Technologie FPGA vyniká kombinací výkonnosti, flexibility, ceny a spotřeby. Tento tým je zastřešený zájmovým sdružením CESNET a spolupracují v něm mladí vědci z Fakulty informačních technologií VUT v Brně, Fakulty informačních technologií ČVUT v Praze a Fakulty informatiky Masarykovy univerzity v Brně pod vedením svých starších a zkušenějších kolegů. Tým úspěšně vytvořil celou řadu technologií, které na trhu jako komplexní produkty nabízí partnerská společnost INVEA-TECH [2], mezi jejíž zákazníky patří největší český internetový vyhledávač Seznam.cz, poskytovatel internetového připojení UPC, Ministerstvo obrany České republiky a další významné instituce [13]

V tomto článku budou představeny problémy datových center - především v oblasti monitorování počítačových sítí. Podrobně se podíváme na kartu COMBO-80G, která tvoří podvozek pro tvorbu hardwarově akcelerovaných síťových aplikací. Jednou z těchto aplikací je i HANIC (hardwarově akcelerovaná síťová karta) [3], nad kterou je postavený komplexní monitorovací systém FlowMon [4], který je založen na technologiích Netflow v9 a IPFIX. Statistiky Netflow a IPFIX lze použít např. pro detekci anomálií na síti [14]. Dnešním trendem je tyto detekce zautomatizovat a provádět v reálném čase. Přesně tyto požadavky si klade za cíl splnit systém Nemea (Network Measurements Analysis) [5], který se taktéž intenzivně vyvíjí v rámci sdružení CESNET. Dále si ukážeme, jakým způsobem lze monitorovat až šestnáct 10Gb linek v jediném serveru U2 (tzv. High-density řešení až pro 160Gb/s!). Nakonec trochu nahlédneme do budoucnosti a zjistíme na co se můžeme v oblasti monitorování sítí těšit.

Datová centra čelí celé řadě problémů. Začíná to výběrem geograficky i politicky vhodné lokality, pokračuje vysokými náklady na výstavbu a končí ještě vyššími náklady na provoz a údržbu. Právě na provoz a údržbu se nejvíce zaměříme. Velkým problémem je vysoká spotřeba elekrické energie - měřících nástrojů nevyjímaje. S nepříliš velkou nadsázkou lze říci, že každý Watt spotřebované energie se platí třikrát: 1. nákup samotné elekrické energie, 2. zajištění chlazení a 3. místo v datovém centru (energeticky náročnější zařízení bývají i větší). Dalším zásadním problémem je zajištění nepřetržité dostupnosti a vysoké bezpečnosti služeb. Musíme řešit nákup výkonnějšího hardware, protože ten starý již nedostačuje. Hardware nebude nikdy 100% spolehlivý, a proto musí být systém odolný vůči poruchám - lze relativně snadno řešit přidáním dalšího hardware (redundance). Vzhledem k tomu, že datová centra poskytují služby online, je naprosto nezbytné zajistit spolehlivou a bezpečnou síťovou konektivitu. Uvedeme si jedno přirovnání: Datové centrum je město a síťová konektivita je dálnice, která do města vede. Stejně jako na dálnici, je nutné při komunikaci po síti dodržovat určitá pravidla k zajištění spokojenosti všech zúčastněných. Funkci policie plní v počítačových sítích firewall, který může určitý síťový provoz zahodit. V této analogii potřebuje ještě firewall nějaká pravidla “silničního” provozu a “zařízení pro sledování provozu”. Jako pravidla

síťového provozu můžeme považovat standardy RFC (doufejme, že autoři pracují lépe než český parlament) a jako zařízení pro sledování provozu síťové sondy (dnes nejčastěji sondy generující statistiky NetFlow v9 a IPFIX). Pro zajištění bezpečnosti a spolehlivosti síťového provozu tedy potřebujeme kvalitní statistiky ze síťových sond a kvalitní detekční algoritmy.

I když vědci ze sdružení CESNET již osedlali technologii 100Gb Ethernetu [7], stále se v datových centrech hojně využívá technologie 10Gb Ethernetu - například český internetový vyhledávač Seznam.cz má ve svých datových centrech pět 10G přípojek [15]. Odpovědí lidí z CESNETu je akcelerační karta COMBO-80G, která umí pracovat v režimu 2x40Gb Ethernetu, nebo 8x10Gb Ethernetu.



Akcelerační FPGA karta COMBO-80G [7]

Podívejme se na klíčové vlastnosti:

- FPGA čip Xilinx Virtex-7 690T
- 2×QSFP+ optické rozhraní pro 40Gb Ethernet
- Možnost konfigurace 2x4x10Gb pro realizaci 10Gb Ethernetu
- Rozhraní PCI Express 3.0 x8 s propustností do operační paměti až 64 Gb/s
- 1x144 Mb rychlé statické QDR paměti na 500 MHz
- 16x4 Gb dynamické DDR3 paměti na 1600 MHz
- Sychnronizační vstup z GPS pro generování velmi přesných časových značek

Kdo se o FPGA technologii blíže zajímá, tak ví, že připájením FPGA čipu na desku plošných spojů práce nekončí, nýbrž začíná. Do FPGA čipu je nutné nahrát komplexní firmware, jehož složitost si nezadá s velkým softwarovým systémem. A to je oblast, ve které lidé z CESNETu

excelují. Pro účely monitorování se používá firmware HANIC (Hardware Accelerated Network Interface Card) [9]. Představme si jeho základní vlastnosti:

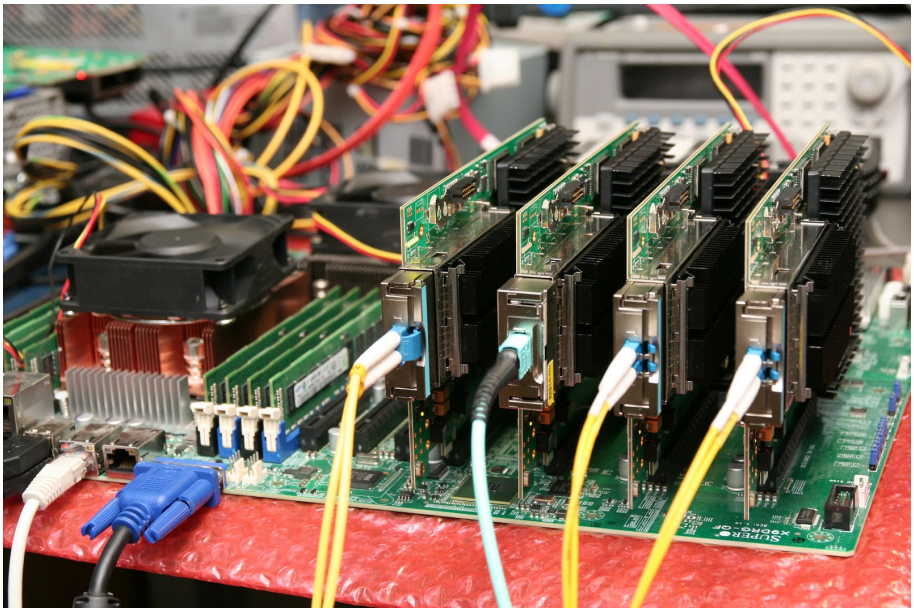
- Odesílání a příjem síťového provozu na 2x40Gb nebo 8x10Gb Ethernetu
- Konstantní výkonnost na všech paketových délkách
- Generování časových značek s nanosekundovou přesností
- Inteligentní distribuce síťového provozu na jednotlivá procesorová jádra (8, 16, ...)
- Výpočet statistik NetFlow v9 a IPFIX
- Funkce jednoduchého firewallu

Ve skutečnosti HANIC statistiky NetFlow v9 ani IPFIX přímo nepočítá. Provádí předzpracování síťového provozu takovým způsobem, aby již samotný výpočet statistik byl co nepřímočarejší (hardwarová akcelerace). Oproti běžně používanému řešení bez hardwarové akcelerace tento způsob šetří výpočetní výkon CPU a tím i elektrickou energii. Použití technologie FPGA je pro síťové aplikace velice vhodné. Zaprvé oproti univerzálním procesorům CPU mají mnohem větší výkon i lepší spotřebu, ale narozdíl od specializovaných hardwarových čipu ASIC [10] nabízí technologie FPGA vysokou flexibilitu, a může tak reagovat na aktuální trendy v počítačových sítích. Výpočet statistik a jejich vizualizaci za použití systému HANIC a karty COMBO-80G provádí komplexní systém FlowMon [4]. Systém poskytuje velmi kvalitní statistiky, včetně časových značek s nanosekundovou přesností, bez nutnosti vzorkování a to na plné rychlosti linky. Přiřazování časových značek k přijatým paketům na síti je velmi důležité z hlediska detekce anomálií a rozhodně není v dnešní době standardem. Myšlenka vzorkování spočívá v tom, že pro výpočet statistik se používá pouze každý n-tý paket (např. 4, 8, ...). Pokud se vrátím k přirovnání s dálnicí, tak to znamená, že policie si chce ulehčit práci a místo toho, aby měřila rychlost vozidla, dodržování bezpečné vzdálenosti, agresivitu jízdy či věnování se řízení, tak se zaměří pouze na rychlost jízdy, protože to je nejjednodušší a veřejnosti bude tvrdit, že rychlost je jedinou metrikou při řízení motorového vozidla.

Dnešním trendem je automatizace detekcí anomálií na síti. Potřebujeme tedy kvalitní detekční algoritmus (např. pro DDoS, Heartbleed, Shellshock a jiné typy útoků) a kvalitní statistiky o síťovém provozu - případně i celý provoz tam, kde je to nutné (zachytávání všeho síťového provozu je prakticky nemožné, proto se používají pouze statistiky, ale selektivně možné je). K automatizované detekci síťového provozu byl vytvořen a stále se vyvíjí systém Nemea [11] zájmového sdružení CESNET, který značně těží z kvalitních statistik poskytnutých systémem FlowMon.

Už jsme si řekli, že problémy datových center jsou spotřeba elektrické energie, místo ve skříních rack a zajištěni síťové konektivity. Systém pro automatizovanou detekci síťových anomálií, popsaný v předchozích odstavcích, se skládá z 2U serveru, akcelerační karty COMBO-80G, firmwaru HANIC, systému FlowMon a Nemea. Takto popsané řešení je schopno monitorovat až osm 10Gb linek. Do serveru je ovšem možno osadit i druhou kartu

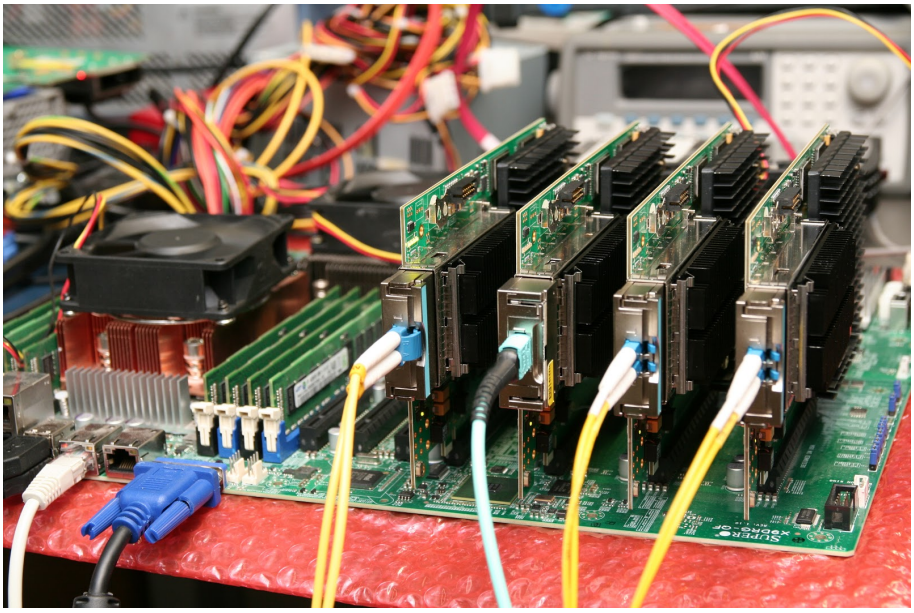
COMBO-80G a tím navýšit schopnosti na monitorování až šestnácti 10Gb linek při typické spotřebě pouze 200W. Pro lepší představu se podívejme na obrázek takového serveru:



High-density řešení 2xCOMBO-80G Chybí!!!

Nazýváme to High-density řešením pro datová centra, kdy jsme schopni poskytnou komplexní službu monitorování pro šestnáct 10Gb linek, při nízké spotřebě elektrické energie a to vše pouze v 1 serveru 2U. Navíc časem, až bude běžnější technologie 40Gb Ethernetu, není nutno měnit jakýkoliv hardware - toto řešení lze snadno upravit (změna firmware v FPGA a software) na monitorovací systém pro 4x40Gb Ethernet. Této flexibility umožnuje docílit technologie FPGA, při použití technologie ASIC, kterou používají jiní konkurenční výrobci, by tato záměna nebyla možná.

V tomto článku jsme si představili High-density řešení pro monitorování šestnácti linek 10Gb Ethernetu. Toto řešení je vhodné pro datová centra, protože má nízkou spotřebu, nabízí komplexní službu monitorování až šestnácti 10Gb linek a navíc se celé poskládá do 1 serveru 2U. Dnes se nejčastěji používá technologie 10Gb Ethernetu, ale v blízké budoucnosti se přejde na technologii 100Gb Ethernetu (40Gb Ethernet se pravděpodobně přeskočí). Vědci ze sdružení CESNET jsou si toho vědomi, a proto již představili akcelerační kartu COMBO-100G [12] a systém HANIC na ní postavený. Dokonce se již experimentuje s High-density řešením pro monitorování 4x100Gb nebo 40x10Gb Ethernetu. Podívejte se sami:



High-density řešení 4xCOMBO-100G

[1] The Project Liberouter. Online
https://www.liberouter.org/
[2] INVEA-TECH - High-Speed Networking and FPGA Solutions. Online
https://www.invea.com/
[3] HANIC. Online
https://www.liberouter.org/technologies/hanic/
[4] FlowMon. Online
https://www.invea.com/cs/products/flowmon
[5] Nemea | Liberouter / Cesnet TMC group
https://www.liberouter.org/technologies/nemea/
[6] Request for Comments – Wikipedie
http://cs.wikipedia.org/wiki/Request_for_Comments
[7] COMBO-100G. Online
https://www.liberouter.org/combo-100g/
[8] FlowMon – Případová studie Seznam.cz, a.s.
https://www.invea.com/data/success-stories/seznam.cz_cz.pdf
[9] HANIC. Online
https://www.liberouter.org/technologies/hanic/
[10] Application-specific integrated circuit - Wikipedia, the free encyclopedia
http://en.wikipedia.org/wiki/Application-specific_integrated_circuit
[11] Nemea | Liberouter / Cesnet TMC group
https://www.liberouter.org/technologies/nemea/
[12] COMBO-100G. Online
https://www.liberouter.org/combo-100g/

[13] References - INVEA-TECH
https://www.invea.com/en/company/references
[14] Network Management Case Study
http://www.cisco.com/web/about/ciscoitatwork/network_systems/network_data_monitoring_and_reporting_web.html
[15] FlowMon - Případová studie Seznam.cz, a.s.
https://www.invea.com/data/success-stories/seznam.cz_cz.pdf

1) Problemy datovych center
a) misto v racku
b) spotreba
c) monitorovani at uz vykonu ci bezpecnosti
2) Nase flexibilni reseni a jak je to reseni dobre a vhodne pro datova centra.

* 320G v 2U serveru: Důkladný monitoring s úsporou nákladů a energie
  high-density pro data centra (FB80G)
  - velký počet portů v 2U krabici - úspora místa, el energie
  - využití FPGA technologie, která je green
  - ASIC řešení se nedá upravovat, není flexibilní, FPGA může reagovat na aktuální trendy z bezpečnosti, atd. Úspora nákladu za HW
  - variabilita portů 2x40Gb nebo 8x10Gb na kartu, úspora investic
  - předzpracování v HW, počítání statistik na komoditním železu, HANIC