

Projekt routeru IPv6

Jiří Novotný, ÚVT MU

1 Úvod

Projekt routeru IPv6 koordinovaný ÚVT MU je zařazen do strategického projektu CESNETu *Implementace IPv6 v síti CESNET 2* jako dílčí úkol *Vývoj hardware pro podporu směrování IPv6* a byl vybrán jako nosný projekt CESNETu při vstupu do konsorcia 6NET, které koordinuje nasazení IPv6 v rámci Evropské Unie.

IPv6 je nová verze Internetového Protokolu, která nahradí stávající verzi IPv4. Od zavedení IPv6 se na jedné straně očekává vyřešení stávajících problémů verze IPv4 (malý adresní prostor, praktická neexistence hierarchického adresování) a na druhé straně rozšíření o další služby (zejména v oblasti Quality of Service a IPSec). V současné době je paradoxně největší překážkou nasazení IPv6 existence stávajících sítí na bázi IPv4. Není možné najednou vyměnit programové vybavení na několika desítkách milionů počítačů. Proto bude přechod z IPv4 na IPv6 postupný a po určitou (zřejmě dosti dlouhou) dobu budou používány obě verze společně. To se projeví ve zvýšených nárocích na routery, které budou muset být schopny nejen zpracovat obě verze IP, ale zabezpečit i jejich vzájemnou komunikaci.

Použití routerů na bázi počítačů PC v Brněnské Akademické Počítačové Síti (BAPS) má dlouhou tradici. Prvním routerem, kterým byl budovaný zárodek BAPS připojen v roce 1992 do Internetu, bylo PC s operačním systémem DOS a hardwarovým akcelerátorem SRP vyvinutým na Ústavu výpočetní techniky Masarykovy univerzity. Dlouholeté zkušenosti s nasazením PC routerů v BAPS jsou velmi dobré. Spolehlivost provozu jak po hardwarové, tak po softwarové stránce je plně srovnatelná s komerčními produkty při podstatně příznivější ceně. Funkční možnosti komerčních i PC routerů jsou minimálně srovnatelné. Velmi důležitou vlastností PC routerů je snadné testování nových služeb, na rozdíl od komerčních routerů, kde je vždy třeba počkat, jestli daný výrobce bude testovanou službu podporovat.

Hlavními technickými omezeními pro širší nasazení PC routerů je jejich poněkud složitější konfigurace a do jisté míry i jejich výkon, který je u špičkových PC routerů, podle měření prováděných v rámci výzkumného projektu CESNETu, srovnatelný s routery střední třídy.

Úkolem předkládaného projektu je navrhnout PC router IPv6 a současně vyřešit obě zmíněná technická omezení. Pro jednodušší konfiguraci a údržbu routeru vyvineme konfigurační prostředí podobné prostředím známým z komerčních routerů (CISCO, Juniper). Zvýšení výkonu PC routeru chceme dosáhnout vývojem hardwarového akcelerátoru na bázi PCI karty, který bude obsahovat několik gigabytových portů a jeho výkon bude v oblasti jednotek Gb/s.

2 Konfigurační prostředí PC IPv6 routeru

Současné PC routery používané v BAPS jsou založeny na operačním systému NetBSD (BSD Unix). Konfigurace se provádí editací souborů v adresáři /etc. Struktura konfiguračních souborů se vzájemně značně liší. Při přechodu na novou verzi operačního systému, případně i jiný operační systém (FreeBSD, Linux) se může struktura i uložení těchto souborů dále měnit, což způsobuje nepřehlednost konfigurace.

Naším úkolem je vypracovat jednotící prostředí, které by uvedené nedostatky eliminovalo. Protože chceme podporovat větší množství platform, vybrali jsme jako toto prostředí jazyk XML, ve kterém bude konfigurace uložena. Jazyk XML je pro běžnou práci s routerem poměrně komplikovaný a proto je nutné vytvořit nad ním uživatelsky přívětivé nadstavby.

V projektu budeme implementovat CLI (Command Line Interface), Webové rozhraní a SNMP. Pro strukturu CLI jsme se rozhodli jako vzor použít routery Juniper. Jejich CLI je dobře zpracované a převod do XML poměrně jednoduchý. Webové rozhraní je určeno méně zkušeným uživatelům a předpokládáme, že alespoň v první fázi bude umožňovat konfiguraci nutný pro vzdálenou správu routerů pomocí nadstavbového software.

Konfigurace zapsaná v některém z příkazových prostředí bude převedena do jazyka XML a z něj

budou generovány potřebné konfigurační soubory pro daný operační systém. Tato poněkud komplikovaná struktura konfiguračního prostředí je výhodná při přechodu na jinou verzi operačního systému, kdy stačí dopsat převodník mezi XML a použitým soubor v /etc. Při doplnění převodníků pro komerční routery (CISCO, Juniper, ...) je možné konfiguraci routeru uchovávat v jednotném formátu, což usnadní vzájemnou nahraditelnost zařízení různých výrobců.

Nadstavbou nad tímto konfiguračním prostředím bude systém, který umožní uchovávat konfiguraci všech routerů v síti v jednotném formátu.

3 Hardwarový akcelerátor

3.1 Návrh architektury PC IPv6 routeru

Vzhledem k vývoji PC počítačů a jejich současnému výkonu lze na této platformě zkonstruovat router s výkonem v řádu stovek Mb/s. Úzkým místem je propustnost PCI sběrnice a rychlost odezvy procesoru na přerušení.

Propustnost sběrnice u běžných PC počítačů (šířka sběrnice 32 bitů, frekvence 33 MHz) je 1 Gb/s. Paket prochází přes sběrnici dvakrát (vstupní a výstupní port) a z toho vychází teoretická propustnost 500 Mb/s. Jistým zlepšením je použití výkonnější sběrnice (64 bitů, 66 MHz), kde lze dosáhnout teoretické propustnosti při routování 2 Gb/s, ale ani to nemusí být v praxi dostatečné.

Druhým problémem je rychlost odezvy procesoru na přerušení, která u výkonnějších strojů dosahuje okolo 80 000 přerušení/sekundu. Rychlost odezvy procesoru je kritická při zpracování velmi krátkých paketů.

Přestože uvedený výkon odpovídá zhruba výkonu routeru CISCO 7500 a je pro mnoho aplikací dostatečný, chtěli bychom zvýšit propustnost PC routeru do oblasti jednotek až desítek Gb/s a to i v oblasti krátkých paketů.

Řešením je rozdělit funkci routeru na části. První bude provádět obslužné, poměrně složité a paměťově náročné operace (výpočet routovacích cest, konfigurace, statistika) s nižšími požadavky na rychlost. Druhá část bude vykonávat vlastní přepínání a filtraci paketů. Zde je požadována

zejména vysoká rychlost, ostatní nároky nejsou tak vysoké.

Pro první část lze s výhodou použít PC architektury, kapacita paměti i procesoru je pro tento účel dostatečná. Pro druhou část navrhne a vyvineme hardwarový PCI akcelerátor, který bude provádět autonomní přepínání paketů.

Vzájemná komunikace přes PCI se tak omezí na inicializaci, přenos konfiguračních a statických dat, případně řešení výjimek (kdy akcelerátor nemá informace potřebné ke korektnímu směrování paketu).

3.2 Hardwarový PCI akcelerátor

Při vývoji hardware je potřeba vyvážit výkon, složitost, rozšiřitelnost, dostupnost komponent, cenu vývoje a cenu produkce daného zařízení. Na základě podrobné analýzy těchto aspektů je potřeba zvolit nejlépe vyhovující architekturu a provést vlastní vývoj. Návrh a vývoj v hardwarové oblasti je na rozdíl od software poměrně dlouhodobá záležitost, opravy chyb v návrhu jsou podstatně složitější.

Velkou změnu při vývoji přinesla myšlenka programovatelného hardware, kdy na čipu integrovaného obvodu je umístěno velké množství poměrně jednoduchých buněk s programovatelnou logickou sítí. Činnost obvodu je pak dána konfigurací jednotlivých buněk a jejich vzájemným propojením. Výhodou je možnost změny funkce hardware na hotové desce se zaletovanými integrovanými obvody pouhou změnou konfiguračních dat. Nevýhodou je poněkud nižší rychlost a vyšší cena v porovnání s „klasickým hardware“. Tyto obvody byly dříve používány většinou pro prototypové verze a kusovou výrobu. V poslední době výrazně klesla cena i u velmi výkonných obvodů spolu se zvýšením jejich rychlosti a proto jsou stále častěji používány pro malosériovou výrobu a výrobu středně velkých sérií.

Při návrhu hardwarového akcelerátoru jsme se rozhodli použít výhod programovatelného hardware. Srdcem akcelerátoru bude hradlové pole FPGA (Field Programmable Gate Array) firmy XILINX, doplněné paměťmi (DRAM, SRAM, CAM), obvody rozhraní a nezbytnou bižuterií (obvody zdrojů a hodin). Celý akcelerátor bude modulární s množstvím rozšiřujících konektorů.

Vývoj v oblasti routovacích protokolů je natolik bouřlivý, že v současné době je velmi obtížné zahrnout všechny funkce v době návrhu hardware. Koncepce akcelérátoru s programovatelným hardware nám umožní pružně reagovat na změny ve vývoji routovacích protokolů, navrhovat a testovat nové funkce.

3.3 Verifikace

Vzhledem k tendenci zvyšující se složitosti navrhovaných integrovaných obvodů a hardwarových systémů obecně, je pro návrháře stále obtížnějším problémem vyvarování se chyb vzniklých nedodržením specifikace. Týmy návrhářů bývají proto doplňovány skupinou verifikátorů.

Standardní metody verifikace zaručují důvěryhodnost systémů úspěšnou simulací navrhovaného hardware na velkém množství testovacích případů. Oproti tomu metody formální verifikace využívají matematických prostředků k analýze celého stavového prostoru navrhovaného systému vzhledem k požadované specifikaci, což odpovídá otestování všech možných situací, v nichž se systém může ocitnout. Tedy zatímco simulace je neúplná metoda a nezaručuje jistotu, formální verifikace eliminuje nejistotu v podstatně vyšší míře.

Formální verifikace je také součástí projektu PC routeru.

4 Závěr

Návrh IPv6 routeru na bázi PC je složitý a dlouhodobý úkol, který vyžaduje spolupráci odborníků v mnoha oblastech. Celý tým pracující na projektu má přes třicet lidí z CESNETu, Masarykovy univerzity, Vysokého učení technického v Brně, doplněného zahraničními specialisty.

Výsledkem vývoje bude výkonný IPv6 router, s velkou škálou funkcí, který bude vhodný k nasazení do akademických sítí s ohledem na další výzkum a vývoj internetových protokolů. □