

MPT AND MPTCP MECHANISMS IN DATA NETWORKS

Petr Sejkora

Master Degree Programme (2), FEEC BUT

E-mail: xsejko00@stud.feec.vutbr.cz

Supervised by: Jaroslav Koton

E-mail: koton@feec.vutbr.cz

Abstract: This work deals with multipath transmission in the data networks to speed up transmission over networks with limited transmission speed. Work compares the characteristics, efficiency and response to changes in individual transmission paths. Work is specifically dedicated to MPT and MPTCP mechanisms.

Keywords: MPT, MPTCP, multipath data transport

1. ÚVOD

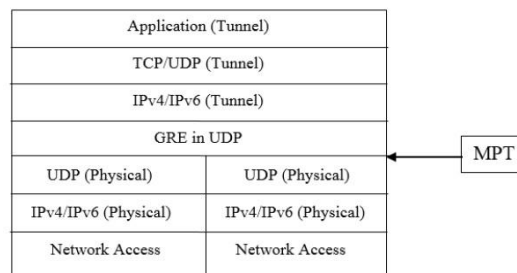
V současné době dochází k připojování nepřeborného množství zařízení k datovým sítím. S tím, jak se zařízení rozšiřují v počtu a schopnostech, se rozšiřuje i nabídka služeb, které jsou skrze zařízení poskytovány. Nároky na kapacitu a spolehlivost datových sítí se tedy zvyšují.

Vývoj probíhá i v rámci datových sítí, které tato zařízení propojují do vyšších celků. Na úrovni koncových zákazníků zpravidla jedna linka stačí. Na úrovni ISP, větších podniků a poskytovatelů služeb už ne. Architektura sítí musí mít také redundanci k zajištění provozu služby i v případě selhání některého prvku. Vznikají tak složitější celky s více cestami.

Pokud máme více cest, přichází na řadu otázka, jak je efektivně využít. Balancování zátěže se věnuje tato práce a klade si za cíl porovnat možnosti, efektivitu a chování vybraných mechanismů sloužících k tomuto účelu.

2. KNIHOVNA MPT

Knihovna vznikla jako projekt pod vedením dr. Béla Almási na Fakultě informatiky University of Debrecen v Maďarsku [1]. Klade si za cíl vytvořit možnost přenosu více cestami na úrovni síťové vrstvy. Využívá IETF RFC s názvem "GRE in UDP" [4], který je zatím ve formě návrhu a rozšiřuje jej o možnost přenosu dat více cestami. Pokud chce aplikace komunikovat s protistranou přes MPT, musí využít předem nakonfigurovaného tunelu s druhou stranou. Odešle požadavek na komunikaci s IP adresou dostupnou na druhém konci tunelu a předá nižší vrstvě. Změna nastává až na linkové vrstvě, kdy se rámec nepředává fyzické vrstvě, ale vstupuje do GRE tunelu. Jeho provoz zajišťuje právě MPT a rozděluje jej do více, až dvaceti, toků datagramů UDP. Ty jsou dále zpracovány standardním způsobem.

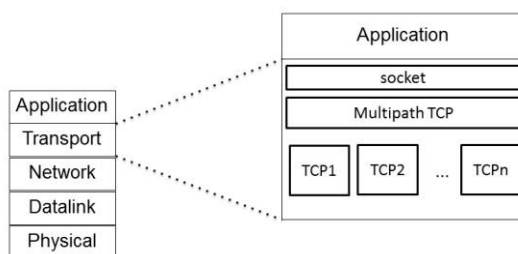


Obrázek 1: Architektura MPT [1]

Aplikace MPT serveru musí být spuštěna na obou stranách tunelu. Může se jednat o koncové stanice nebo o směrovače spojující celé sítě. Konfigurace nabízí obě možnosti. Provádí se formou zápisu dat do souborů, které si MPT server načte při spuštění. Po spuštění je možné posílat příkazy na vzdálený server pomocí klienta. Servery si mezi sebou posílají keepalive zprávy. Knihovna vytváří tunel s IP adresou na obou koncích a dokáže tedy přenášet jak UDP, tak TCP, případně IPv6 skrze IPv4. Vzhledem k tomu, že je tento tunel provozován na UDP protokolu, není zaručeno doručení odeslaných dat. Je tedy obtížné řízení toku dat cestami, které nejsou stejné nebo vykazují odchylky ve spolehlivosti, propustnosti nebo zpoždění. Nelze tedy očekávat bezproblémový přenos. Protokol umožňuje použití vyrovnávací paměti před výstupem z tunelu (pro seřazení paketů, pokud záleží na pořadí), ale když dojde ke ztrátě během přenosu, tak následky neřeší. V případě přenosu TCP se chybějící segment vyšle znovu nad úroveň tunelu, ale UDP datagram je ztracen. V konfiguraci je dále možné nastavit počet a váhu cest, sledování funkčnosti dané cesty nebo autentizaci řídicích zpráv [1].

3. MULTIPATH TCP

Vývoj tohoto protokolu začal jako magisterská práce Dr. Sébastiena Barré na Université catholique de Louvain v Belgii [2]. Momentálně vývoj vede Dr. Christoph Paasch. MPTCP se snaží využívat přenosu více cestami pro přenos jednoho TCP spojení za účelem lepšího využití zdrojů, vyšší propustnosti a plynulejší reakce na výpadek [3]. Aplikace vytvoří požadavek na TCP spojení na danou IP adresu. Spojení se navazuje použitím 3-way handshake. Strana, která vytváří spojení, odešle paket s příznakem SYN a MP_CAPABLE, ve kterém je i šifrovací klíč na pozdější použití. Pokud má protistrana o spojení zájem odešle též SYN+ACK+MP_CAPABLE se svým klíčem. Navazující strana potvrdí ACK a tím je vytvořena relace MPTCP s prozatím jedním spojením. Pokud má jedna ze stran více adres k dispozici, tak se pokusí navázat s protistranou další spojení zprávou s příznakem MP_JOIN a identifikátorem, ke které relaci MPTCP se chce připojit. Protokol tedy musí rozlišovat sekvenční čísla jednotlivých subrelací i nadřazené relace a spravovat rozdělení toků. Spojení se tedy navazuje na požádání a veškerá konfigurace se provádí dynamicky dle aktuálních podmínek v síti.



Obrázek 2: Architektura MPTCP [2]

Vývojáři si kladli za cíl schopnost reálného nasazení protokolu, aby mohl fungovat v reálném prostředí stávajících sítí a Internetu. Nasadit novou technologii na všechny prvky Internetu je nemožné a některé směrovače a firewally zahazují provoz, který nedokáží identifikovat. Jako jedinou možnost tedy shledali postavit protokol tak, aby fungoval na již existující architektuře, bez zásahu. Protokol je specifikován v RFC 6824 [3].

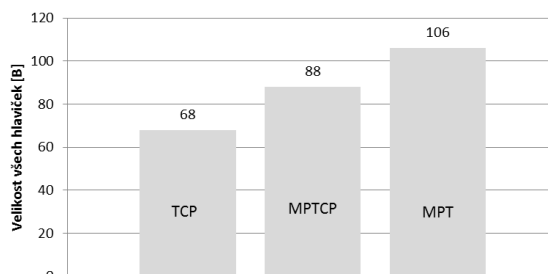
4. TESTOVACÍ SESTAVA

Při testování mechanismů TCP a MPTCP a pro jejich porovnání se základním protokolem TCP byla vyhodnocována jejich režie a schopnost využít více paralelních cest z pohledu celkové dosažené přenosové rychlosti na aplikační vrstvě a času potřebného k přenosu daného objemu dat. Měření probíhalo ve virtuálních strojích VirtualBox v OS Ubuntu. Hostitelský počítač měl dvě fyzická jádra. Každý stroj měl 1Gb/s síťovou kartu, přes kterou mohl komunikovat s druhým OS. Generování zátěže zajišťoval program IPerf, který na dané sestavě byl schopný vygenerovat

TCP provoz až 21,2 Gb/s. Omezení šířky pásma každé cesty na 10 Mb/s zajišťoval třetí počítač se dvěma síťovými kartami a programem WANem.

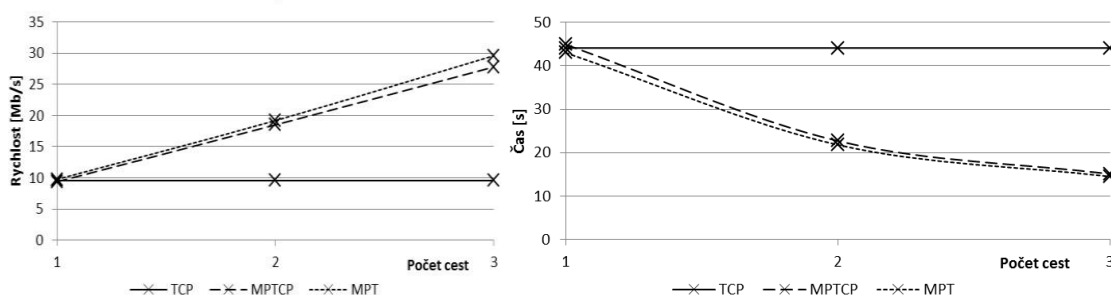
4.1. VELIKOST HLAVIČEK

Protokoly vnáší do přenosu kromě zpoždění a nároků na výkon také další data do hlaviček paketů. Problém nastává hlavně u MPT, kde v našem případě často docházelo k překročení MTU linky, fragmentaci a následně další reži. Proto je nutné věnovat pozornost konfiguraci MTU tunelu.



Obrázek 3: Režie protokolů

4.2. NAMĚŘENÉ VÝSLEDKY



Obrázek 4: (a) Rychlost přenosu

(b) Čas potřebný pro přenos 50MB dat

Rychlost rostla s počtem linek téměř lineárně a linky byly vytěžovány na 100%. Režie samotných protokolů co do množství dat je zanedbatelná (viz. Obr. 3). V tomto ohledu se jejich nasazení určitě vyplatí. Jiná situace může nastat, pokud se začnou měnit vlastnosti jednotlivých linek. MPTCP bude na změny reagovat, MPT nikoli. Když byla rychlost linek limitována zahazováním nadbytečných paketů, propustnost protokolu MPT na třech linkách byla ještě nižší než samostatný tok TCP.

5. ZÁVĚR

V článku byly diskutovány mechanismy MPT a MPTCP, které dokáží využít více cest pro spojení mezi stanicemi.

Nutnou podmínkou provozování protokolu MPT je prvotní manuální konfigurace koncových stanic, kdy se tak předpokládá, že definovaná spojení jsou navázána staticky a protokol sám jen omezeně reaguje na dynamiku provozu na definovaných linkách.

Výhody protokolu MPT jsou ve schopnosti přenosu jak UDP, tak TCP spojení přes až 20 různých linek. Je možné ovládat server z druhé strany příkazovým kanálem. Protokol je možné použít i jako spoj mezi sítěmi.

Naopak protokol MPTCP je plně automatizovaný a reaguje dynamicky na změny parametrů sítě. Pracuje na spolehlivém protokolu TCP, který mu umožňuje okamžitě reagovat na změny. Spojení

se navazuje, když je potřeba (on demand) v závislosti na počtu a stavu linek a po uzavření přenášeného spojení se ukončuje.

REFERENCE

- [1] B. Almási, The MPT-GRE Multipath Communication library, available from <http://irh.inf.unideb.hu/~almasi/new/index.php/projektek/19-mpt-library>
- [2] C. Paasch, S. Barre, et al., Multipath TCP in the Linux Kernel, available from <http://www.multipath-tcp.org>
- [3] RFC6824: TCP Extensions for Multipath Operation with Multiple Addresses, available from <https://tools.ietf.org/html/rfc6824>
- [4] RFC “GRE in UDP Encapsulation”, available from <https://tools.ietf.org/html/draft-ietf-tsvwg-gre-in-udp-encap-11>