

P2P Protokoly pro přenos souborů*

Andrei Yakuta

Slovenská technická univerzita v Bratislavě
Fakulta informatiky a informačních technologií
xyakuta@stuba.sk

15. prosince 2023

Abstrakt

Tento článek se zabývá Peer-to-Peer (P2P) protokoly pro přenos souborů, které umožňují sdílení dat bez centrálního serveru. P2P protokoly jako jsou BitTorrent, eDonkey a Gnutella, jsou zkoumány s ohledem na jejich architekturu, vyhledávání a sdílení souborů, a bezpečnostní opatření. Dále článek porovnává P2P a tradiční client-server modely, hodnotí různé implementace P2P protokolů a diskutuje možný vývoj P2P technologií a jejich potenciální dopad na budoucí digitální aplikace a služby. Závěrem poskytuje ucelený pohled na aktuální stav a perspektivy P2P technologií v oblasti přenosu souborů.

1 Úvod

P2P (Peer-to-Peer) protokoly pro přenos souborů představují způsob sdílení dat mezi uzly ve stejné síti bez nutnosti centrálního serveru. Tento způsob sdílení dat umožňuje výrazné snížení nákladů na infrastrukturu, zatímco zajišťuje rychlý a spolehlivý přenos velkých souborů. Protokoly Peer-to-Peer jsou v dnešní době velmi oblíbené, protože umožňují decentralizaci od serverů (což v dnešní době zabraňuje “zájemcům” mazat soubory, a tím podporuje svobodu slova).

V mém článku se zaměřím na analýzu různých P2P protokolů, jako jsou BitTorrent, eDonkey a Gnutella. Hlavním cílem je porozumění architektuře těchto protokolů, mechanismům vyhledávání a sdílení souborů, a způsobům, jak tyto protokoly řeší problémy s bezpečností a soukromím. Porovnám také P2P protokoly s tradičními client-server modely, abych diskutoval o výhodách a nevýhodách obou přístupů. Důležitou součástí mé analýzy bude hodnocení různých implementací P2P protokolů a jejich dopad na výkon, škálovatelnost a odolnost proti chybám. Dále se pokusím předvídat možný vývoj P2P technologií a zvážit jejich potenciální dopad na budoucí aplikace a služby v digitálním světě.

2 Co je Peer-to-Peer?

Pro počítačové systémy Peer-to-Peer (P2P) jsou charakteristická konsensuální spojení mezi “peery” (rovnocenní účastníci sítě). To znamená, že počítače v síti mohou vzá-

*Semestrální projekt v předmětu Metody inženýrské práce, ak. rok 2023/24, vedení: Ing. Richard Marko, PhD.

jemně komunikovat, aniž by měly procházet přes samostatný serverový počítač [3]. Z tohoto důvodu se při přechodu z Client-Server sítě na Peerto-Peer snižuje zatížení serveru, což je jedna z výhod používání protokolu P2P. Navíc se zvyšuje rychlost přenosu dat a propustnost sítě.

Pokud jde o P2P síť, je charakteristické, že každý účastník sítě může plnit funkci serveru i klienta. Při přenosu souborů prostřednictvím P2P protokolů to znamená, že každý účastník sítě, který stáhl část přenášeného souboru (funkce klienta), ji může předat ostatním účastníkům (funkce serveru). Přitom se vůbec (nebo téměř vůbec) nepoužívá samostatný server.

Jednou z klíčových vlastností tohoto typu sítě pro přenos souborů je, že síť skoro není závislá na jediném serveru, tj. síť se nezhroutí, pokud ji někteří členové opustí, na rozdíl od sítě Client-Server, kde při výpadku serveru dojde k výpadku celé sítě.

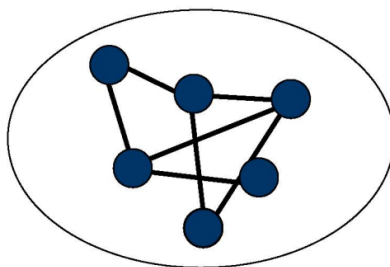
Za zmínku stojí také to, že díky decentralizaci sítě P2P téměř není možné zastavit šíření jakýchkoli souborů v síti, protože všichni uživatelé v síti jsou servery. Jinými slovy, na rozdíl od sítě Client-Server, kde by distribuce souborů skončila po vypnutí serveru, by museli být všichni členové sítě vyřazeni.

3 P2P protokoly pro přenos souborů

Ačkoli nejzákladnějším požadavkem na síť P2P je decentralizace od serveru, některé protokoly používají servery ke zjišťování informací o souborech nebo o uživateli, od nichž lze určitý soubor převzít. Na základě toho lze protokoly rozdělit do dvou různých skupin podle jejich architektury: čistá P2P a P2P zprostředkovaná serverem [11].

3.1 Čisté P2P protokoly

K této skupině patří protokoly jako Gnutella a Freenet. Hlavní výhodou čistých P2P protokolů je, že síť, na rozdíl od P2P protokolů zprostředkovaných serverem, vůbec není závislá na serveru, protože žádný nemá. Ale kvůli té nezávislosti musí dynamicky objevovat další partnery v síti, což způsobuje problémy, jako je vyšší spotřeba síťových prostředků a nízká škálovatelnost [11].

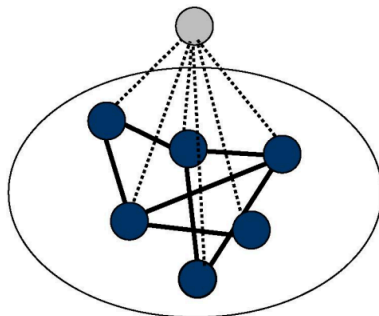


Obrázek 1: Architektura čisté P2P sítě. Zdroj [10]

3.2 P2P protokoly zprostředkované serverem

K této skupině patří protokoly jako eDonkey, Napster, a BitTorrent. Jak bylo dříve zmíněno, jednou z nevýhod této skupiny P2P protokolů je, že síť je závislá na serveru.

I když se síť při výpadku serveru nezhroutí, P2P aplikace nebude schopna najít další peers [11].



Obrázek 2: Architektura P2P sítě zprostředkované serverem. Zdroj [10]

4 Srovnání P2P a tradičních client-server modelů

P2P a tradiční client-server modely představují dva různé přístupy k organizaci a řízení síťové komunikace a sdílení dat. Přestože mají oba modely své silné stránky a jsou vhodné pro různé účely, odlišují se v několika klíčových aspektech.

Aspekt	P2P	Client-Server
Decentralizace	Ano	Ne
Škálovatelnost	Vysoká	Omezená
Náklady	Rozložené	Centralizované
Bezpečnost	Variabilní	Vyšší
Rychlost	Vysoká	Závislá na serveru
Anonymita	Ano	Ne

Tabulka 1: Srovnání P2P a client-server modelů

4.1 Decentralizace versus centralizace

P2P model je charakterizován decentralizací, kde každý uzel (peer) v síti může fungovat jak jako klient, tak jako server, což podporuje distribuci dat a zajišťuje robustnost sítě při výpadku jednotlivých uzlů. Na druhou stranu, client-server model je založen na centralizované architektuře, kde server poskytuje služby klientům a spravuje přístup k datům. Výpadky serveru v client-server modelu mohou vést k výpadkům služby, zatímco v P2P modelu může síť nadále fungovat, dokud existuje dostatečný počet peerů.

4.2 Škálovatelnost

P2P sítě mají potenciál ke škálování, protože noví uživatelé mohou přispívat svými zdroji (např. bandwidth, úložným prostorem) k celkové kapacitě sítě. Naopak, client-server modely mohou čelit omezením v škálovatelnosti, pokud server nedokáže zvládnout rostoucí počet klientů.

4.3 Náklady na infrastrukturu

V P2P modelu jsou náklady na infrastrukturu distribuovány mezi všechny účastníky sítě, což může vést k nižším celkovým nákladům ve srovnání s client-server modely, kde je potřeba investovat do silných a spolehlivých serverů, aby bylo možné zvládnout požadavky klientů [9].

4.4 Bezpečnost

Client-server modely umožňují lepší kontrolu a správu dat a služeb, což může vést k vyšší úrovni bezpečnosti a spolehlivosti. V P2P modelu může být obtížnější zajistit bezpečnost a integritu dat, zejména v otevřených P2P sítích, kde může dojít k většímu riziku neoprávněného přístupu a šíření škodlivého softwaru [9].

Avšak v P2P sítích může decentralizovaná struktura ztížit vnějším pozorovatelům, jako jsou útočníci nebo autority, zjistit, kde jsou data uložena a kdo je sdílí, což může uživatelům poskytnout vyšší úroveň anonymity. Naopak, v tradičních client-server modelech je sledovatelnost jednodušší, protože veškerá komunikace prochází centrálním serverem, který může být snadno monitorován a analyzován.

4.5 Rychlost a výkon

Výkon a rychlost přenosu dat může být ve velkých P2P sítích výrazně vyšší, protože data mohou být přenášena paralelně od mnoha peerů. Na druhou stranu, výkon client-server modelů může být omezen kapacitou serveru, zejména při vysoké zátěži.

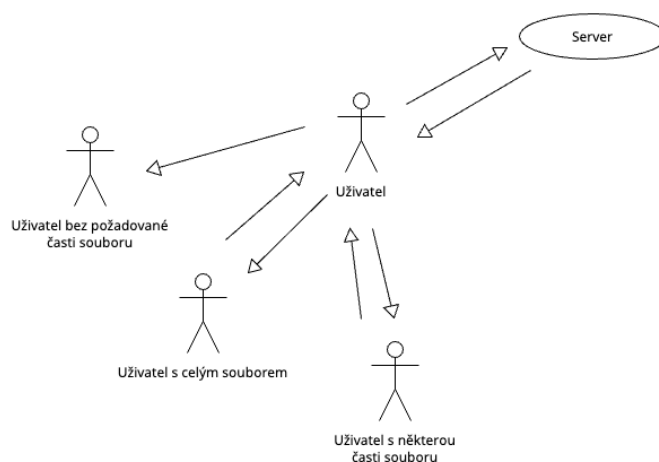
5 Hodnocení a dopad P2P protokolů

5.1 BitTorrent protokol

BitTorrent je příkladem P2P sítě s centralizovanou strukturou (zprostředkovanou serverem). Centrálním uzlem v této síti je tracker (server), který obsahuje informace o seznamu uzlů připojených k síti a službách poskytovaných jednotlivými uzly (např. seznam souborů dostupných ke stažení z daného uzlu). Pro stažení požadovaného souboru uzel odesílá trackeru požadavek (pomocí HTTP protokolu) obsahující identifikátor požadovaného souboru, který se obvykle získává z torrent souboru (viz dále) [4]. Tracker odpovídá na tento požadavek seznamem uzlů, ve kterých je požadovaný soubor k dispozici [13]. Server nadále pomáhá uzlům vzájemně komunikovat, ačkoli poslední verze softwaru vyžadují server pouze v inicializační fázi (což umožňuje větší decentralizaci od serveru) [7].

Specifikem BitTorrentu je pojem torrent, který definuje relaci přenosu části souboru do množiny peerů. Uživatel se obvykle připojí k existujícímu torrentu stažením .torrent souboru, který obsahuje metainformace o stahovaném souboru a IP adresu trackeru torrentu [4, 8].

Podle protokolu BitTorrent se soubory přenášejí po částech, přičemž každý klient tyto části stahuje současně s tím, jak je předává ostatním klientům. Tím se snižuje zatížení a závislost na každém zdrojovém klientovi [7]. Při dostatečně velkém počtu peerů se tím také výrazně zrychlí stahování souborů, což je velká výhoda tohoto protokolu při stahování velkých souborů [2].



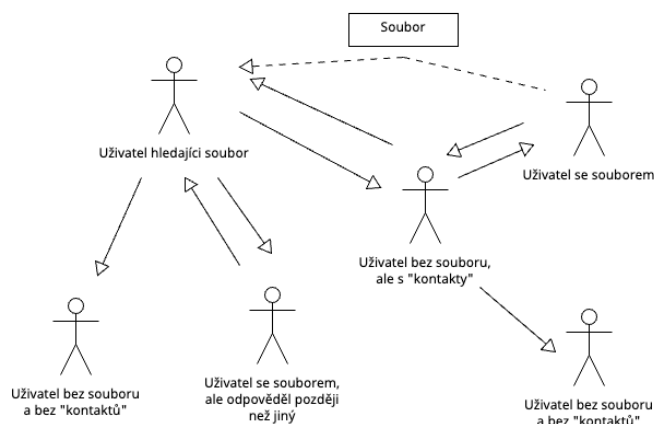
Obrázek 3: Architektura BitTorrent sítě

5.2 Gnutella protokol

Gnutella je příkladem P2P sítě s decentralizovanou strukturou (čistě P2P), která nevyžaduje centrální servery pro indexaci nebo vyhledávání souborů. Na rozdíl od BitTorrentu, který využívá tracker pro koordinaci sdílení souborů, Gnutella funguje na plně decentralizovaném modelu. Každý uzel v Gnutella síti může fungovat jak jako klient, tak jako server, což umožňuje přímé sdílení souborů mezi uživateli. Tato architektura zajišťuje, že síť je odolnější proti výpadkům a obtížněji cenzurovatelná.

V Gnutella protokolu, když uživatel hledá soubor, jeho dotaz se přeposílá od uzlu k uzlu pomocí mechanismu zvaného "flooding". Tento proces spočívá v rozesílání dotazu všem sousedním uzlům, které pak tento dotaz dále přeposílají, dokud není nalezen odpovídající soubor [5]. Skutečné stahování souboru probíhá pomocí HTTP portokolu s uzlem, který první informoval, že soubor má [1, 12]. Ačkoliv tento způsob je efektivní pro nalezení souborů v malé síti, může způsobovat vysokou síťovou zátěž v rozsáhlejších sítích [14].

Navzdory výzvám spojeným s decentralizací, jako jsou efektivita vyhledávání a bezpečnostní rizika, Gnutella představuje důležitý model v P2P technologiích. Jeho přístup k distribuci a sdílení souborů zůstává významný pro svou schopnost podporovat svobodný a neomezený přenos informací mezi uživateli. Avšak decentralizovaná povaha sítě ztěžuje implementaci bezpečnostních kontrol, a proto je Gnutella často vnímána jako méně bezpečná než jiné P2P sítě, jako je BitTorrent.



Obrázek 4: Architektura Gnutella sítě

6 Budoucnost P2P technologií

S rostoucí digitalizací a vývojem technologií se očekává, že P2P technologie budou hrát klíčovou roli v budoucích inovacích. Jedním z nejperspektivnějších směrů je vývoj nových protokolů, které spojují výhody P2P a tradičních client-server modelů. Tyto hybridní protokoly by mohly nabídnout vyšší efektivitu, lepší škálovatelnost a větší bezpečnost.

P2P sítě by tak mohly být lépe integrovány do cloudových služeb, což by umožnilo lepší správu a distribuci dat. Kromě toho se očekává, že P2P technologie budou hrát významnou roli v oblastech jako decentralizované úložiště dat, správa velkých dat a vývoj decentralizovaných aplikací (DApps), které jsou stále více populární v blockchainovém prostředí. Tyto inovace by mohly výrazně ovlivnit způsob, jakým lidé sdílejí a spravují digitální obsah, a otevřít cestu pro nové formy digitální interakce a kolaborace. [6]

7 Závěr

V tomto článku byla prozkoumána dynamika a vliv Peer-to-Peer protokolů v kontextu současného digitálního světa. Porovnáním s tradičními client-server modely jsou odhaleny jedinečné výhody P2P přístupu, zejména v decentralizaci, škálovatelnosti a efektivitě přenosu dat. Prostřednictvím případových studií protokolů jako BitTorrent a Gnutella je ukázáno, jak různé P2P systémy přistupují k přenosu souborů a jak řeší výzvy, jako jsou bezpečnost, vyhledávání a distribuce dat.

Význam P2P technologií se neustále zvyšuje v době, kdy digitalizace a požadavky na ukládání dat rychle rostou. Vzhledem k jejich potenciálu pro rozvoj decentralizovaných aplikací a integraci s blockchainovými technologiemi, se očekává, že P2P protokoly budou hrát klíčovou roli v budoucích technologických inovacích. Budoucnost P2P technologií může zahrnovat nové hybridní protokoly, které kombinují nejlepší aspekty P2P a client-server modelů, čímž poskytnou silnější, bezpečnější a efektivnější systémy pro sdílení a správu digitálního obsahu.

Závěrem lze říci, že P2P technologie nejenže představují významnou část současného digitálního ekosystému, ale budou také hrát zásadní roli v jeho budoucím vývoji.

Ať už jde o podporu svobody projevu, efektivní distribuci dat, nebo podporu nových inovací, P2P protokoly jsou klíčovým prvkem pro budoucí digitální krajinu.

Reference

- [1] The gnutella protocol specification v0.4. URL https://courses.cs.washington.edu/courses/cse522/05au/gnutella_protocol_0.4.pdf.
- [2] C. Barakat and I. Pratt. P2p and overlay. In *Passive and Active Network Measurement: 5th International Workshop, PAM 2004, Antibes Juan-les-Pins, France, April 19-20, 2004, Proceedings*, volume 3015, pages 1--33, Antibes, 2004. Springer. URL <https://books.google.sk/books?id=LZD1BwAAQBAJ&pg=PA1#v=onepage&q&f=false>.
- [3] M. Bauwens, V. Kostakis, and A. Pazaitis. *Peer to Peer: The Commons Manifesto*. University of Westminster Press, London, 2019. URL <https://library.oapen.org/handle/20.500.12657/25288>.
- [4] A. Chokkalingam and F. Riyaz. Bittorrent protocol specification v 1.0, 2004. URL https://www.academia.edu/1032175/BitTorrent_Protocol_Specification_V_1_0.
- [5] A. J. Howe. Napster and gnutella: a comparison of two popular peer-to-peer protocols. *Universidade de Victoria*, 11, 2000. URL https://members.tripod.com/ahowe_ca/pdf/napsternutella.pdf.
- [6] P. Kiseembe and W. Jeberson. Future of peer-to-peer technology with the rise of cloud computing. *International Journal of Peer to Peer Networks (IJPP2P)*, 8(2/3):45--54, 2017. URL https://www.researchgate.net/profile/Phillip-Kiseembe/publication/319579735_Future_of_Peer-To-Peer_Technology_with_the_Rise_of_Cloud_Computing/links/5a68409f0f7e9b76ea8fdbb9/Future-of-Peer-To-Peer-Technology-with-the-Rise-of-Cloud-Computing.pdf.
- [7] D. Lande. P2p — по секрету всему свету... О пиринговых сетях [p2p - důvěrně světu... o sítích peer-to-peer]. *Сети и бизнес [Sítě a podnikání]*, 39(2):104--110, Květen 2008. URL <http://dwl.kiev.ua/art/p2p/p2p-end.pdf>.
- [8] A. Legout, G. Urvoy-Keller, and P. Michiardi. Understanding bittorrent: An experimental perspective, 11 2005. URL <https://inria.hal.science/inria-00000156/document>.
- [9] K. Leibnitz, T. Hoßfeld, N. Wakamiya, and M. Murata. Peer-to-peer vs. client/server: Reliability and efficiency of a content distribution service. In L. Mason, T. Drwiega, and J. Yan, editors, *Managing Traffic Performance in Converged Networks*, pages 1161--1172, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg. URL https://link.springer.com/chapter/10.1007/978-3-540-72990-7_99.

- [10] P. Leong, C. Miao, B. K. Lim, J. W. Lim, C. Chen, N. Dianna, and W. B. Toh. Agent mediated peer-to-peer mobile service-oriented architecture. In *2007 Inaugural IEEE-IES Digital EcoSystems and Technologies Conference*, pages 414--419, 2007. URL <https://ieeexplore.ieee.org/document/4233742>.
- [11] S. M. Lui and S. H. Kwok. Interoperability of peer-to-peer file sharing protocols. *ACM SIGecom Exchanges*, pages 25--33, Sprn 2002. URL https://www.researchgate.net/publication/220553583_Interoperability_of_peer-to-peer_file_sharing_protocols.
- [12] M. Portmann, P. Sookavatana, S. Ardon, and A. Seneviratne. The cost of peer discovery and searching in the gnutella peer-to-peer file sharing protocol. In *Proceedings. Ninth IEEE International Conference on Networks, ICON 2001.*, pages 263--268, 2001. URL <https://ieeexplore.ieee.org/document/962351>.
- [13] G. I. Radčenko. *Распределенные вычислительные системы [Distribuované výpočetní systémy]*, chapter Технологии одноранговых сетей [Technologie Peer-to-Peer sítí], pages 117--130. Министерство образования и науки Российской Федерации, Южно-Уральский государственный университет, Кафедра системного программирования [Ministerstvo školství a vědy Ruské federace, Jihouralská státní univerzita, katedra systémového programování], Čeljabinsk, 2012. URL http://ww.w.glebradchenko.ru/doc/Radchenko_Distributed_Computer_Systems.pdf.
- [14] M. Ripeanu. Peer-to-peer architecture case study: Gnutella network. In *Proceedings First International Conference on Peer-to-Peer Computing*, pages 99--100, 2001. URL <https://ieeexplore.ieee.org/document/990433>.