

Davor Perišić
0036377380

WHITE PAPER
Peer-2-Peer Networks

Sadržaj

1. Uvod.....	3
2. Što je to P2P.....	4
3. Zašto P2P.....	5
4. Kako radi internet P2P.....	6
5. P2P sustavi.....	6
5.1 Raspodjeljeno procesiranje.....	6
5.2 Razmjena datoteka.....	7
5.3 Kolaboracija.....	8
5.4 P2P platforme.....	9
6. Problemi i težnje današnjih P2P sustava.....	9

1. Uvod

Do prije desetak godina cijeli internet je počivao na jednom modelu veze: velika računala bila su konstantno povezana i imala stalne IP adrese. Pojavom weba, osobna računala počela su ostvarivati povremeni pristup internetu, te se javila potreba za dinamičkim dodjeljivanjem IP adrese pri svakom spajanju. Ovakav pristup onemogućava osobnim računalima da se ponašaju kao poslužitelji podataka ili mrežno orjentiranih aplikacija.

Nekoliko godina PC je tretiran kao glup ali skup mrežni klijent i to je dobro funkcioniralo. Međutim, PC posjeduje mnogo veći potencijal, pa zašto to ne iskoristiti i na mreži? Na internetu su u svakom trenutku spojeni milijuni PCa, što predstavlja ogromnu količinu procesorske moći i prostora za pohranu.

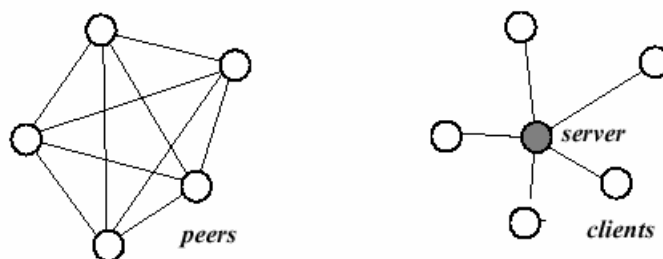
Tu na scenu dolaze Peer-2-Peer mreže (P2P) ili mreže međusobno povezanih *peerova*. *Peer* je običan PC normalnog internet korisnika, najčešće modemsom vezom spojen na internet, bez stalne IP adrese. P2P veza nije novi pojam jer, uostalom, mrežni poslužitelji oduvijek komuniciraju P2P vezom. Ono što nove P2P mreže omogućuju je međusobno povezivanje velikog broja PCa trenutno spojenih na internet, mimo DNS servera, bez stalnih IP adresa i uz djelomičnu ili potpuno neovisnost o centralnim poslužiteljima. Na taj način, dobiva se jedan decentralizirani dinamički mrežni sustav, cijenom održavanja i resursima raspodjeljen na tisuće PCa diljem interneta. Ovakav mrežni sustav može se iskoristiti na nebrojene načine, od razmjene podataka do raspodjele procesorskog posla.

1. Što je to P2P?

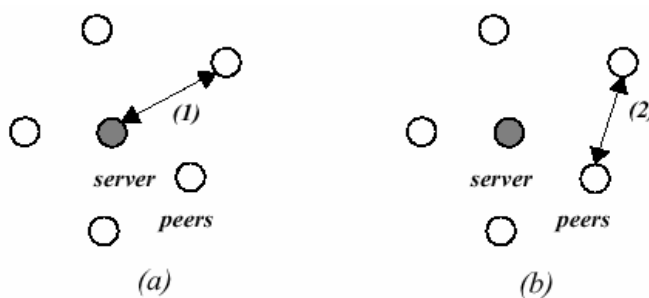
U svom najizvornijem obliku, P2P označava razliku u odnosu na klijent/poslužitelj model mreže koji je dominirao mrežama pred-internet razdoblja, a dominira i strukturom samog interneta danas. Ideja klijent/poslužitelj modela mreže je u tome da se resursi (npr. datoteke) nalaze na poslužitelju, a klijenti do tih resursa mogu doći šaljući zahtjev poslužitelju. Poslužitelj u tom slučaju mora obrađivati zahtjeve i posluživati sve klijente koji se na njega spajaju.

S druge strane, P2P mreže omogućavaju klijentima da dijele resurse međusobno. Svaki klijent je tada *peer*, tj. ima jednaka prava i mogućnosti i može započinjati komunikaciju. U nekim slučajevima, P2P komunikacija je izvedena tako što se svakom komunikacijskom čvoru pridjele mogućnosti i klijenta i servera.

Internet P2P mreže nude usluge slične onima što su ih u počecima razvoja mreža nudili *MS Windows for Workgroups*, *Sun's Network File System* ili *File Transfer Protokol (FTP)*, iako s mnogo sofisticiranijim pretraživanjem i kretanjem i sve to na javnoj internet mreži. Važnu predhodnicu današnjim P2P mrežama predstavlja tzv. «anonimni FTP», tj. FTP veza koja dopušta razmjenu datoteka bez prethodne identifikacije klijenta.



P2Pvs.client/server



Hybrid Peer-to-Peer Model.

(a) upit serveru o adresi *peera* (b) komunikacija među *peerovima*

3. Zašto P2P?

Općenito, internet ima tri važna svojstva: informacije, propustnost i računalna sredstva, od kojih su svi nedovoljno iskorišteni, djelomično zbog tradicionalnog klijent/server modela.

Kao prvo, niti jedna tražilica (search engine) ili portal ne može locirati i katalogizirati stalno povećavajuću količinu informacija na webu u nekom razumnom vremenu. Pored toga, ogromna količina podataka ne može se dohvatiti trenutno korištenim metodama pretraživanja. Prema istraživanjima, na svijetu se proizvede oko 2×10^{18} byta informacija svake godine, ali se samo oko 3×10^{12} byta objavi. Drugim riječima, za svaki megabyte informacija samo se jedan byte objavi. Jedna od najvećih tražilica, Google, tvrdi da pretražuje samo oko 1.3×10^8 web stranica. Prema tome, traženje korisne informacije u realnom vremenu je sve teže.

Drugo, unatoč velikim količinama novih kablova koji se ugrađuju, nova propustnost ima malu korisnost od toga ako svi idu istovremeno na npr. Yahoo po neke sadržaje. Naprotiv, vruće točke postaju još toplije dok hladne linije ostaju hladne. To je razlog zašto korisnici još uvijek imaju osjećaj zagušenosti na internetu dok se propustnost povećala za faktor 10^6 od 1975, udvostručujući se svakih 16 mjeseci.

Treće, novi procesori i uređaji za pohranu podataka nastavljaju rušiti rekorde u brzini i kapacitetu, podupirući sve snažnije konačne uređaje kroz cijelu mrežu. Međutim, obrada podataka se nastavlja akumulirati oko centara informacija, koja moraju povećati količinu posla u velikoj brzini, čime se povećava potreban prostor i potrošak energije.

P2P arhitektura se odnosi na široki pojas tehnologije koja izrazito povećava iskoristivost informacija, propustnosti i računalnih sredstava na internetu. U većini slučajeva, te P2P tehnologije prihvaćaju mrežno orjentirani model koji niti isključuje niti inherentno ovisi o centraliziranim kontrolnim točkama. Osim povećanja performansi u traženju informacija, njihovom dostavljanju i obradi, ovakav model može poboljšati i ukupnu pouzdanost računalnih sustava.

4. Kako radi internet P2P?

Peer je entitet koji "razgovara" protokolima koji su od njega traženi. Ovo je slično Internetu, gdje je Internet čvor entitet koji "razgovara" nizom IP protokola. Kao takav, *peer* može imati oblik procesora, procesa, uređaja ili korisnika.

Pošto ne postoji centralni poslužitelj, klijenti (*peerovi*) sami obavljaju i poslužiteljske i klijentske poslove pomoću posebne programske podrške koju ima instaliranu svaki *peer*. Budući da se ove mreže koriste u razne svrhe, od razmjene datoteka do raspodjele procesorskog posla, postoje i razni namjenski programi koji ih pokreću (npr. Gnutella za razmjenu datoteka ili SETI@Home za raspodjeljeno procesiranje). *Peer* PC se poveže na drugi *Peer* PC koji se nalazi u dotičnoj mreži (najčešće pomoću IP adresa objavljenih na webu), te preko njega na slijedeći i tako na sve ostale dok se potpuno ne poveže na mrežu.

5. P2P sustavi

Software praktički čini samu P2P mrežu. P2P software se može u grubo podijeliti na tri sloja. Na dnu se nalazi središnji sloj koji se bavi uspostavom peer-ova i komunikacijskim upravljanjem kao na primjer routanjem. U sredini se nalazi uslužni sloj koji se bavi konceptima višeg ranga kao indeksiranje, traženje i izmjenom podataka. Na vrhu se nalazi sloj aplikacije kao što je e-mailanje i spremanje podataka. Neke osobine, kao što je sigurnost, nalaze se u sva tri sloja i svugdje unutar P2P sustava, iako u različitim oblicima ovisno o lokaciji u software-skoj arhitekturi.

Danas se uglavnom javljaju četiri različita tipa P2P sustava:

- raspodjeljeno procesiranje
- razmjena datoteka
- kolaborirajući sustavi
- P2P platforme

5.1 Raspodjeljeno procesiranje

Raspodjeljeno procesiranje se vrlo uspješno ostvaruje na P2P mrežama. Koncept se sastoji u tome da se iskoristi neaktivnost PCa spojenih na internet, da se procesorski

zahtjevan posao podijeli na veliki broj malih, manje zahtjevnih poslova koji se izvršavaju svaki na svom *peer* PCu. Svaki PC opremljen je softwareom koji koristi vrijeme neaktivnosti PCa za obradu dotičnih podataka, te komunicira sa upravljačkim poslužiteljem. Središnji upravljački poslužitelj vrši raspodjelu poslova, te obrađuje i spaja rezultate pojedine obrade. Iako koristi upravljački poslužitelj, raspodjeljeno procesiranje se ipak ubraja u P2P sustave zbog autonomije koju *peerovi* imaju.

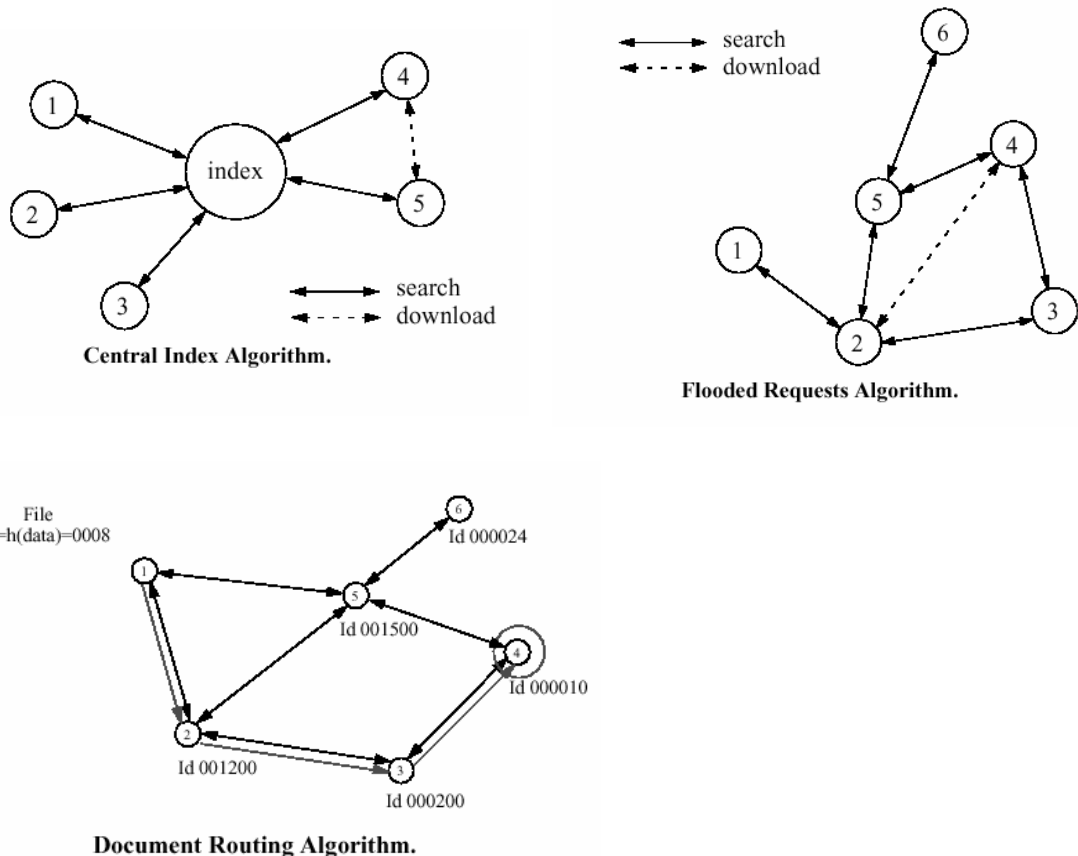
Najveće ograničenje u primjeni raspodjeljenog procesiranja je zahtjev za poslom s mogućnošću raspodjele u nezavisne male djelove, bez potrebe za komunikacijom među *peerovima*. Najčešća upotreba raspodjeljenog procesiranja na P2P mrežama je istraživanje ljudskog genoma, istraživanja na području lijekova protiv raka, financijske i biotehnološke primjene. Jedan od najpoznatijih projekata ovog tipa P2P sustava je SETI@home koji se bavi potragom za vanzemaljskom inteligencijom.

5.2 Razmjena datoteka

Pohrana i razmjena datoteka je možda najuspješnija primjena P2P tehnologije. Datoteke multimedijskog sadržaja zauzimaju mnogo prostora i iziskuju veliku propusnost pri prijenosu putem interneta. Raspodjeljeni sustavi pohrane bazirani na P2P tehnologiji omogućavaju:

- *Razmjenu datoteka*. Koristeći redundanciju, ovakvi sustavi nude potencijalno neograničen prostor za pohranu. Datoteke se spremaju na neke čvorove u P2P mreži, ali je dostupna svakom *peeru*. Da bi *peer* mogao doći do tražene datoteke, mora samo znati referencu na dotičnu datoteku. Primjeri ovakvih sustava su Freenet, Gnutella i Kazaa.
- *Vrlo dostupno sigurnosno pohranjivanje*. Postojanje virtualnog prostora za pohranu i multiplikacija kritičnih datoteka povećava njihovu dostupnost.
- *Anonimnost*. Neki P2P sustavi kao što je Publius omogućavaju da objavljeni dokumenti zadrže anonimnost njihovih autora i izdavača, a učine ih dostupnima korisnicima mreže.
- *Upravljivost*. P2P sustavi omogućavaju lagan i brz pronalazak i dohvat traženih datoteka raspodjeljujući podatke po *cache*-ovima raspoređenim po rubovima mreže. Položaj podataka nije poznat *peeru* koji zahtjeva podatak i može biti raspodjeljen duž prijenosnog puta, te se u potpunosti gubi princip *hostanja*. Tu se mogu javiti problemi pouzdanosti prijenosa cjelovite datoteke, kao i odgovornosti za ponudene podatke.

Glavna tehnička pitanja u sustavima razmjene datoteka su propusnost mreže, sigurnost i pretraživanje. Danas su zastupljena tri modela: model centralnog direktorija (Napster), model *flooded* zahtjeva (Gnutella) i model usmjeravanja dokumenata (FreeNet).



5.3 Kolaboracija

Kolaborativne P2P aplikacije ciljaju na to da omoguće suradnju na nivou aplikacije među korisnicima mreže. *Peer to peer* svojstvo P2P sustava čini ga pogodnim za suradnju na razini korisnika. Upotreba ovakvih aplikacija seže od sustava za direktnu razmjenu poruka (*instant messaging* i *chat*), *online* igara, do dijeljenih aplikacija u poslovnoj, edukacijskoj i kućnoj upotrebi.

Kolaborativne aplikacije su općenito bazirane na *eventima* (dogadajima ili promjenama). *Peerovi* oforme grupu i započnu s nekim poslom. Kada se neki event javi u jednom *peeru* (npr. *peer* pošalje poruku), on se šalje svima na grupi. Na aplikacijskoj razini, sučelje svakog korisnika reagira na promjenu.

Postoje brojni problemi koji otežavaju implementaciju ovakvih kolaborativnih sustava. Kao i u ostalim P2P sustavima, tu je problem lociranja ostalih *peerova*. Negdje se to rješava centralnom listom trenutno aktivnih korisnika, a negdje pomoću IP adresa (MS NetMeeting). Možda najveći izazov u implementaciji kolaborativnih sustava predstavlja komunikacija u stvarnom vremenu. U nekim slučajevima je vrlo bitno da svi korisnici dobiju obavjest o eventu u isto vrijeme. Nažalost, to uglavnom ne ovisi o samom P2P sustavu već o mrežnoj infrastrukturi koja se nalazi ispod njega. Ovaj problem osobito je važan u okruženju *real-time* mrežnog igranja (npr. DOOM).

5.4 P2P platforme

Operativni sustavi postaju sve manje značajni kao okolina aplikacije. *Middleware* rješenja kao što su Java Virtual Machines ili Web browseri dominantno su okruženje za aplikacije. P2P platforme predstavljaju okruženje korisnicima i uslugama spojenim u P2P mreže. P2P platforme, čak više nego ostali P2P sustavi, imaju podršku za primarne P2P komponente: imenovanje, komunikacija, sigurnost i agregaciju informacija. Imaju minimalnu ovisnost o operativnom sustavu. Poznate P2P platforme su .NET i JXTA.

6. Problemi i težnje današnjih P2P sustava

Interoperabilnost. Mnogi P2P sustavi su građeni za pružanje jedne vrste usluga. Na primjer, Napster omogućava razmjenu muzike a Gnutella omogućava razmjenu generičkih podataka. S obzirom na razlike ovih usluga i zbog nedostatka opće prihvaćene P2P infrastrukture, svaki proizvođač P2P software-a stvara međusobno nekompatibilne sustave. To znači da svaki proizvođač stvara svoju P2P korisničku zajednicu. Da bi peer sudjelovao u različitim zajednicama organiziranim od drugih P2P implementacija, peer mora podržavati višestruke implementacije, od kojih je svaka za određeni P2P sustav.

Neovisnost platformi. Mnogi P2P sustavi danas nude svoje usluge kroz niz API-a (Application Programming Interface) koji su dostavljeni na određenom operacijskom sustavu koristeći određeni mrežni protokol. Na primjer, jedan sustav može nuditi niz C++ API-a, sa sustavom koji inicijalno radi na Windows-ima preko TCP/IP protokola, dok drugi sustav nudi kombinaciju C i Java API-a koji rade na različitim UNIX

sistemima preko TCP/IP ali isto zahtjevajući HTTP. P2P projektant je prema tome prisiljen birati koji niz API-a da koristi tj. prisiljen je odabrati koju grupu P2P kupaca želi privući. Ako se žele ponuditi jednake usluge objema grupama, moraju se razviti jednake usluge za dvije platforme.

Sveprisutnost. Mnogi P2P sustavi pretežno biraju Windows-e kao njihovu razvojnu platformu. Razlog za to je da se želi ciljati na najveću instaliranu bazu i najbrži put do zarade. Ovakav pristup je očigledno kratkovidan pošto P2P ne stoji za PC-To-PC. Najvjerojatnije će se najveća raširenost P2P tehnologije javiti na dva kraja hardwear-skog spektra, veliki sustavi unutar kompanija i korisnički orijentirani mali sustavi.