



marka tomasik

Wymiana plików przez Internet staje się coraz łatwiejsza

Wytrwały pobierze więcej

Napster nie istnieje, ale pobieranie plików MP3, AVI i innych przez Sieć kwitnie. Nowe pomysły programistów sprawiają, że kolejne aplikacje P2P są coraz skuteczniejsze.

Michał Olszowski

Wymiana plików – głównie MP3 – jest niezmiennie jednym z najgorętszych tematów internetowych. Wszyscy pamiętamy pojawienie się Napstera, jego szybką karierę i głośny upadek. Gdy program Fanninga był na szczycie, pojawiały się inne aplikacje służące do tego samego celu. I choć wytwórnie muzyczne oraz filmowe chcą zlikwidować proceder wymiany plików, to technologie peer-to-peer rozwijają się dynamicznie. O niezwykłości Napstera zadecydował nowatorski pomysł, by użytkownicy pobierali zbiory nawzajem od siebie. Kolejne sieci P2P – Gnutella, Direct Connect czy też KaZaA – powstawały i powstają cały czas. Zastosowane w nich rozwiązania nie są kopią technologii Napstera. W głowach programistów cały czas rodzą się nowe pomysły, które wymianę plików MP3, AVI i innych czynią szybszą, bardziej niezawodną i odporną na działania organizacji zrzeszającej wytwórnie płytowe (RIAA).

Więcej swobody

Napster swoje powodzenie zawdzięczał m.in. wysokiej skuteczności wyszukiwania. Było to możliwe dzięki wykorzystaniu centralnego serwera, który gromadził informacje o zbiorach przechowywanych na dyskach

użytkowników. To samo rozwiązanie przyczyniło się jednak do łatwego zablokowania sieci Napstera. Gdy po wyroku sądu przestał działać główny komputer, internauci stracili możliwość wymiany plików. W wielu nowych sieciach P2P zrezygnowano ze scentralizowanej architektury. Twórcy nowych technologii dążą do uniezależnienia ich od jakichkolwiek „strategiczych” punktów. Daje to użytkownikom większą anonimowość, a także uniemożliwia szybkie zamknięcie sieci.

Jedną z pierwszych rozproszonych sieci była Gnutella – w jej przypadku nie istniały żadne wyróżnione komputery (patrz: CHIP 7/2001, 152). Niestety – brak centralnych punktów sieci spowodował, że skuteczność wyszukiwania plików nie była zbyt wysoka. Pośrednie podejście zastosowano w jednym z najnowszych programów P2P – eDonkey2000. Istnieją tu serwery pośredniczące w wyszukiwaniu plików, ale ich prowadzeniem nie zajmują się twórcy oprogramowania (tak było w przypadku Napstera). Powstała sieć niezależnych serwerów, które są rozmieszczone na całym świecie i nie są od siebie zależne. Schemat działania sieci eDonkey znajduje się w ramce „Nowe rozwiązania”.

Teoretycznie taki centralny punkt może założyć każdy. Aby jednak dobrze spełniał

on swoje zadanie, musimy dysponować szybkim komputerem i łączem o dużej przepustowości. Użytkownik, chcąc znaleźć wybrany plik, łączy się z odpowiednim serwerem – w zależności od tego, czego szuka. Przykładowo: jeśli ktoś chce pobrać polski film, powinien znaleźć krajowy serwer.

Koniec z fałszerstwem

W wielu sieciach P2P użytkownikom zdarza się pobrać plik o zawartości nieodpowiadającej nazwie, czyli tzw. fałszywkę (ang. fake). Twórcy programów do wymiany zbiorów opracowali technologie zapobiegające takim sytuacjom. W eDonkey2000, jako w jednym z pierwszych programów P2P, zastosowano technikę tzw. hashowania (patrz: CHIP 10/2002, 123). Jest ona dość prosta, a polega na utworzeniu na podstawie zawartości zbioru sumy kontrolnej. Liczba ta powstaje jako efekt zastosowania bardzo skomplikowanego wzoru i jeśli w pliku zmienimy choćby jeden bajt, powstanie zupełnie inna wartość.

Opisane rozwiązanie znakomicie nadaje się do znajdowania błędów w zbiorach (pobieramy plik i jego sumę kontrolną, po czym tworzymy ją ze zbioru znajdującego się na naszym dysku. Jeżeli wartości są różne –

The screenshot shows the nmap application window. At the top, there's a menu bar with 'File', 'Server', 'Search', 'Options', and 'Console'. Below the menu bar is a toolbar with icons for file operations, server management, search, options, and console. The main area displays a list of discovered hosts in a table format.

nr	ip	port	user	file	name
294	61.16.116.4	6681	2081	416206	DunagySYS-EM Japanese AI center, Max 400
295	213.171.118.6	6681	457	32309	Stank-Italy Server (www.stank.net) Server V.3
317	211.201.219.0	4661	421	13600	Yuhan School Korea Server (Yuhan School Kor
62	64.168.0.13	4661	990	136127	Q1-4 Drinking (GSJC ADSI, 684/394, PH 115)
11	80.2.230.6	4661	419	52873	EuskalOrkestra PHV 1.5 GHz 2x40 GB V9K

nr	ip	port	user	file	name
1226	213.23.64.14	4661			
1680	80.22.114.12	4661			
1163	217.228.229.6	4661			
136	217.81.251.13	4661			
142	217.227.42.3	4661			
38	172.184.47.18	4661	83	833	T-Q-4 online AMD Thundebolt 1000 MHz - M2
166	196.112.128.22	4661			
590	80.132.5.9	4661			
1088	80.132.7.7	4661			
1146	217.233.20.9	4661			
1200	217.223.125.5	4661			
15	196.166.228.28	4661			

Dzięki podłączeniu się **JEDNOCZEŚNIE DO WIELU SERWERÓW** skuteczność wyszukiwania plików w mDonkey znacznie rośnie.

Technologie wymiany plików

plik jest uszkodzony). Dzięki niemu eDonkey sprawdza także w prosty sposób, czy dwaj użytkownicy udostępniają ten sam plik. Dzięki temu program do identyfikacji zbiorów w mniejszym stopniu wykorzystuje ich nazwy.

Hashowanie daje też inne możliwości. Zostało wykorzystane w technice ICH (Intelligent Corruption Handling – w wolnym tłumaczeniu: „inteligentne radzenie sobie z błędami”). Rozwiązanie to pozwala na zmniejszenie ilości danych, które należy ponownie skopiować w przypadku wystąpienia błędów transmisji. Zwykle, jeżeli plik okaże się uszkodzony, musimy pobrać go od nowa. W większości przypadków jednak niepoprawnych jest tylko kilka lub kilkanaście bajtów. W aplikacji wykorzystującej ICH zbiór jest ponownie pobierany – bajt po bajcie. Nowo skopiowane informacje są dodawane do pliku i ponownie liczona jest suma kontrolna – tak długo, aż okaże się, że zbiór został naprawiony. Oczywiście może się zdarzyć, że błąd wystąpił na końcu pliku i wtedy konieczne jest powtórne pobranie całości. Ale często uszkodzenie ma miejsce wcześniej i w takiej sytuacji technika ICH skraca czas potrzebny na naprawienie zbioru. Programiści zamierzają udoskonalić tę funkcję, tak aby aplikacja była w stanie wysłać zapytania o dodatkowe sumy kontrolne do innych użytkowników sieci i przez to zmniejszyć ilość danych do ponownego pobrania nawet o 95%.

Plik na lince

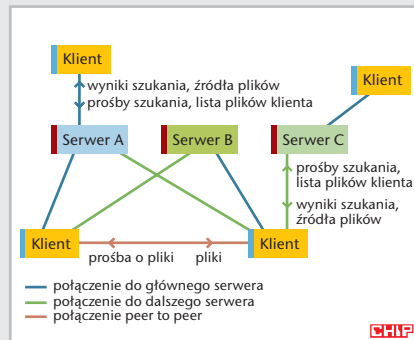
Technika wznawiania przerwanych pobrań (ang. resumimg) jest znana od dawna. Niektóre programy (np. iMesh) wykorzystują ją, aby pobierać plik od kilku osób jednocześnie. Dość znacznie przyspiesza to proces kopiowania. Niestety – dopóki cały zbiór nie znajdzie się na dysku, nie ma mowy o jego udostępnieniu. W eDonkeyu rozwinięto ten pomysł. „Osiołek” dzieli udostępniany plik na części po 9 MB (pliki mniejsze niż 9 MB nie są dzielone). Każdy kawałek może być pobierany od wielu użytkowników, także podłączonych do wielu serwerów. Nowością jest to, że te dziewięciomegabajtowe części, które mamy już na dysku, mogą brać udział w procesie wymiany. Dzięki temu przepływ danych następuje szybciej i sprawniej.

Tworzenie sum kontrolnych przyczyniło się do powstania kolejnego ciekawego rozwiązania, jakim są odsyłacze do plików. Wykorzystując tę technikę, niektóre programy

Nowe rozwiązania

» Architektura sieci eDonkey

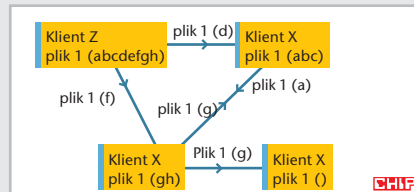
eDonkey2000 wykorzystuje Multisource File Transfer Protocol. Sieć tworzą dwa rodzaje programów: klienty i serwery. Uruchamiając aplikację kliencką, łączymy się z wybranym serwerem, który zawiera informacje o danych znajdujących się na dyskach podłączonych do nich użytkowników. Po otrzymaniu zapytania centralny komputer przeszukuje spis udostępnionych plików i odsyła do użytkownika informacje, kto ma poszukiwane zbiory. Klienci eDonkey mają także opcję szukania rozszerzonego. Jeśli z niej skorzystamy, pytanie wysyłane jest też do innych



serwerów, do których nie jesteśmy bezpośrednio podłączeni.

» Dzielenie plików

Pobieranie pliku z wykorzystaniem technologii podziału zbioru na części sprawia, że dane szybciej rozchodzą się po Internecie. Przykładowo: użytkownik Z ma wszystkie części zbioru 1 (małe litery oznaczają fragmenty danych). Osoby oznaczone literami W, X i Y chcą pobrać plik 1. Jeśli X oraz Y dysponują już niektórymi częściami pobieranego zbioru, to mogą



kopiować je zarówno od użytkownika Z, jak i od siebie nawzajem.

P2P, takie jak eDonkey2000 lub najnowsza wersja klienta KaZaA, modyfikują ustawienia przeglądarki WWW zainstalowanej w naszym systemie i tworzą nowy, obsługiwany przez nią protokół. Żeby dokładnie zrozumieć, jak to działa, najlepiej spojrzeć na rysunek na górze następnej strony. Do wyrażen http, ftp, mailto i kilku innych dodawane jest na przykład ed2k (w przypadku eDonkeya). W odsyłaczu, poza wspomnianym już ciągiem znaków, znajduje się nazwa pliku i jego suma kontrolna. Po kliknięciu odnośnika przeglądarka WWW wywołuje program, który następnie zaczyna wyszukiwać i – jeśli ta operacja zakończy się sukcesem – pobierać wskazany plik.

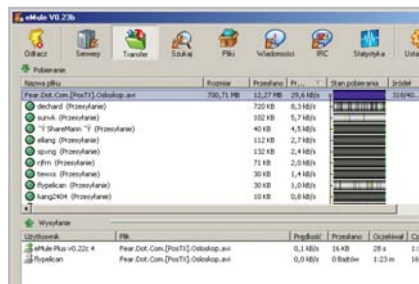
To ułatwienie przyczyniło się do powstania i rozwoju internetowej społeczności

eDonkey. Niemal zaraz po skończeniu prac nad aplikacją i udostępnieniem jej pierwszej pełnej wersji powstały lokalne strony z odnośnikami do plików oraz światowe, komercyjne serwisy, jak np. Sharereactor (www.sharereactor.com). Oprócz pomocy w konfiguracji, ciekawostek i nowinek ze świata „Osiołka” użytkownik znajduje tam odsyłacze do serwerów eDonkeya oraz informacje o nowościach udostępnionych przez grupy specjalizujące się w tworzeniu nielegalnych pakietów oprogramowania (często pozbawionego zabezpieczeń) lub multimediów.

Z opisanych do tej pory technik hashowanie i pakietowa transmisja danych są implementowane w większości nowych programów P2P.

Otwarte głowy

Mimo iż wszyscy pamiętają, jaki los spotkał Napstera, wiele firm próbuje zarabiać na wymianie plików w Internecie, umieszczając w aplikacjach wtyczki reklamowe. Nie chcąc robić konkurencji własnym produktom, firmy nie ujawniają specyfikacji protokołów ani nie udostępniają kodu źródłowego aplikacji. Jednak programiści „rozgrzyżają” wykorzystywane mechanizmy i powstają nowe klienty istniejących sieci. Autorzy kolejnych programów nierzadko wymyślają nowe rozwiązania



Wiele programów P2P potrafi pobierać plik OD WIELU OSÓB JEDNOCZEŚNIE – na przykład eMule.

Technologie wymiany plików

ed2k://|file|joe151win32install_by_counter-strike.de.exe|872278|99c17d9e5bc768f051a61a8ec784d27f|/

i dlatego klony znanych aplikacji często są doskonalsze od pierwowzorów.

eMule, alternatywny klient sieci eDonkey2K, nie powstał na podstawie kodu źródłowego oryginalnego programu, ale sam jest rozpowszechniany na zasadach OpenSource. Działa on w oparciu o tę samą sieć serwerów co eDonkeya2000, wykorzystuje te same odsyłacze do plików i ich sumy kontrolne. Obecnie eMule jest we wczesnym stadium rozwoju, więc niektóre jego funkcje nie działają jeszcze prawidłowo. Co więcej – nie wszystkie opcje eDonkeya są zaimplementowane w jego open-source'owym klonie. Należy jednak mieć nadzieję, że będzie on nadal rozwijany i znajdziemy w nim nowe, ciekawe funkcje. Jednak już obecnie istniejąca wersja zawiera kilka ciekawych pomysłów.

Zmora rozproszonych sieci P2P jest konieczność znalezienia punktu wpięcia do nich – może być nim inny użytkownik (w przypadku Gnutelli) lub serwer (w eDonkeyu). Programy klienckie przechowują informacje o ostatnio aktywnych maszynach, ale są sytuacje, kiedy takie rozwiązanie zawodzi (na przykład gdy długo nie korzystaliśmy z aplikacji). eMule dodatkowo potrafi jednak pobrać listę serwerów z witryn, które je publikują. Dzięki temu program uruchamia się znacznie szybciej. Inną wartościową techniką, o której już pisałem, a która pojawiła się właśnie w eMule, jest Intelligent Corruption Handling.

Jedno z najciekawszych rozwiązań obecnych w otwartym kliencie eDonkey to system „kredytów” przyznawanych użytkownikom. Przeznaczony on jest nie tylko dla klientów eMule'a, lecz dla wszystkich programów sieci eDonkey2000. Internauta podłączający się do komputera współużytkownika sieci jest identyfikowany po sumie kontrolnej oraz dwóch pierwszych

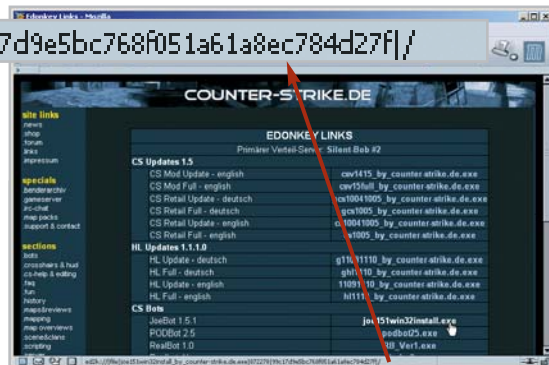
numerach adresu IP i zapamiętywany. Na tej podstawie monitorowane są informacje o pobranych od niego i wysłanych mu danych. Za każdym razem, gdy klient wysłał zapytanie o plik, zebrane wartości będą używane do wyliczenia miejsca w kolejce po dane. Im więcej plików udostępniamy, tym szybciej i więcej zbiorów możemy pobierać. Podobny pomysł był wykorzystywany w nieistniejącym już kliencie P2P – MojoNation. Jego twórca, firma HiveCache, rozpoczął starania o uzyskanie patentów na stosowane tam technologie.

Ostatnia technologia, której celem jest przyspieszenie transferów, to AMUC (Archive Maximal Upload Capacity). Jest to system ustalający parametry wysyłania danych (upload), tak aby ich transfer w tym kierunku był jak największy, a jednocześnie nie przeszkadzało to w pobieraniu zbiorów. Uzyskiwane jest to przez pomiar tzw. czasu ping – jeżeli się on zwiększa, to dane wysyłane są wolniej, a gdy maleje – szybciej.

Jak pingwin z pingwinem

Wiele programów P2P powstaje też w wersjach dla Linuksa. Nierzadko działają one sprawniej i wymieniają dane dużo szybciej niż ich odpowiedniki pracujące w środowisku Windows. Linuksowy klient Direct Connect – DCTC – pod względem funkcjonalności bije swojego okienkowego odpowiednika na głowę. Do najciekawszych rozwiązań DCTC należy GDL (Group DownLoad) – funkcja umożliwiająca jednoczesne pobieranie pliku od wielu użytkowników. DCTC pozwala jednak oszukiwać (np. co do ilości udostępnianych danych), dlatego jest często blokowany przez operatorów tzw. hubów, czyli „serwerów” Direct Connect).

Równie interesującą technologię zastosowano w mDonkey – alternatywnym linuksowym kliencie sieci eDonkey2000. Siła programu tkwi w możliwości łączenia się z wieloma serwerami naraz, podczas gdy okienkowe klienty sieci eDonkey potrafią nawiązać połączenie tylko z jednym serwerem. Oznacza to możliwość znalezienia większej liczby źródeł danych i zwiększenie transferu. Ponadto mDonkey ma bardzo rozbudowane opcje konfiguracji, co pozwala dostosować program do indywidualnych wymagań i zwiększyć szybkości transferu danych. Wiele serwerów



ODSYŁACZE DO UDOSTĘPNIANYCH PLIKÓW sprawiają, że wystarczy jedno kliknięcie, a KaZaA albo eDonkey2000 zaczyna szukać właściwego zbioru.

eDonkey2000 blokuje jednak ten program, ponieważ w niektórych kręgach panuje opinia, że przeciąża on sieć „Osiołka”.

Bez limitu

Upadki Napstera oraz Audiogalaxy zdeterminowały rozwój bardziej zdecentralizowanych technologii peer-to-peer. RIAA wciąż prowadzi krucjatę przeciwko sieciom P2P, należy się więc również spodziewać rozpowszechnienia szyfrowania danych i innych rozwiązań poprawiających prywatność. Sieci P2P, pierwotnie wykorzystywane wyłącznie do udostępniania i kopiowania plików MP3, obecnie służą do wymiany danych wszystkich rodzajów (filmów, programów etc.). Opisane technologie to nie koniec – na pewno pojawią się kolejne rozwiązania usprawniające wymianę plików – by można było pobierać szybciej, więcej i bezpieczniej. ■

INFO

SHAREREACTOR

<http://www.sharereactor.com/>

POLSKIE FORUM eDONKEY2000

<http://www.ed2k.z.pl/>

POLSKIE FORUM OVERNETU

<http://www.overnet.medi.net.pl/>

TECHNOLOGIA P2P

<http://www.openp2p.com/>

eDONKEY2000

<http://www.edonkey2000.com/>

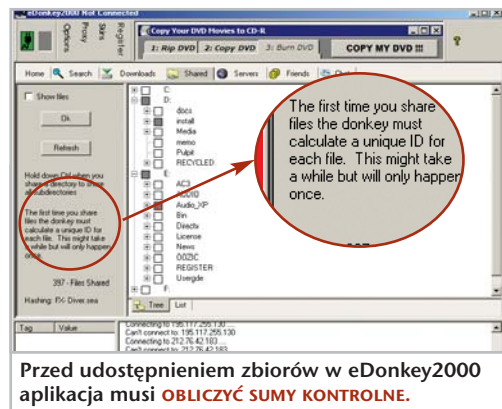
eMULE

<http://www.emule-project.net/>



Na CHIP-CD w dziale Internet i g@dżęty | Wymiana MP3 oraz w sekcji Download | MP3 | Systemy wymiany plików CHIP Online znajdują

się programy Kazaa 2.0.2, iMesh 3.1, eDonkey2000 35.16.61, mDonkey 2.01, eDonkey2000 16.16.59 Lin i eMule 0.24b.



Przed udostępnieniem zbiorów w eDonkey2000 aplikacja musi OBLICZYĆ SUMY KONTROLNE.