

PlanetLab – model budoucího Internetu

Jiří Navrátil, CESNET z.s.p.o.

Bude se Internet podstatně měnit?

V poslední době se v odborných kruzích stále více hovoří o nutnosti změnit základy Internetu, protože s jeho růstem se objevuje stále víc a víc problémů. Nové aplikace, které zajímají miliony lidí, a trvalé zvyšování přípojných kapacit, které umožňuje realizovat přenosy ohromných objemů dat velkými rychlostmi, přináší nejen nové možnosti, ale prohlubují i palčivé problémy, které jsou známy již mnoho let: nedostatečný adresový prostor, přetížení některých fundamentálních služeb, pomalé vyhledávání zdrojů, nespolehlivá stabilita přenosových cest, stále častější pirátské útoky na servery... Jedním z důvodů tohoto stavu je fakt, že mnohé služby a protokoly, které se v Internetu používají, jsou staré desítky let a vzhledem k tomu, že na nich stojí celý Internet, tak se nemohly principiálně změnit, ale jen postupně modifikovat.

PlanetLab <http://www.planet-lab.org> je první „organizace“, která si ve svých cílech vytyčila změnu Internetu. Podívejme se proto, co je to vlastně za organizaci, jaké je její poslání, v jakých směrech se angažuje, co umožňuje a jaké aplikace jsou na ní vyvíjeny a testovány. PlanetLab vznikla v průběhu roku 2002 jako konsorcium několika amerických universit (University of California at Berkeley, Princeton University a University of Washington) [1] a postupem času do ní vstoupily další university z celého světa. Jejimi členy se stala i významná výzkumná pracoviště firem z oblasti IT (jako jsou HP, Intel, France Telecom), organizace zajišťující provoz Internetu v akademických komunitách (jako jsou Internet2-USA, Canarie-Kanada, Cernet-Čína) a další národní pracoviště se širokým výzkumným posláním (INRIA-Francie, GIST-Korea) atd. Konsorcium PlanetLab představuje jednak celosvětovou distribuovanou laboratoř pro výzkum a ověřování nových typů síťových aplikací v planetárním měřítku, současně jde ale i o unikátní síť, samostatnou infrastrukturu, která má 631 uzlů distribuovaných ve všech částech světa. Evropa má v síti

asi 100 uzlů, ale ČR a ani Slovensko v ní zatím zastoupeny nejsou.

Akademické prostředí bylo vždy na čele rozvoje komunikačních systémů a stálo i u zrodu Internetu. Žádná universita a ani žádný stát nebude mít nikdy dostatek prostředků a lidí na to, aby mohl samostatně vytvořit laboratoř, která bude působit v celosvětovém měřítku. To, že se to v projektu PlanetLab podařilo, je velkým krokem vpřed a svědčí to o tom, že PlanetLab má zajímavý a široký program. V minulosti již bylo několik pokusů takové experimentální celosvětové sítě vytvořit, ale ty se většinou zaměřovaly jen na monitorování provozu Internetu (NIMI, Surveyor) a brzy zanikly, protože jejich poslání bylo příliš úzké. Za experiment v tomto směru se dá považovat i propojení gridů. Tam však v konečném řešení půjde o vytvoření produkční sítě, která bude sloužit především pro zajištění dostatečných výpočtových kapacit komunitě uživatelů z oblasti fyziky vysokých energií, pro NASA, chemiky, atd. PlanetLab je něco jiného, je to otevřená laboratoř se širokým posláním pro tvůrce síťových aplikací, tedy prostředí pro implementace návrhů nových komunikačních protokolů, metod distribuovaného zpracování a uchování dat – a jejich testování.

Internet se mění stále

Původní představa a dlouholetý provoz Internetu byly postaveny na konceptu, kdy na jedné straně existuje množina uživatelů (klienti) a na druhé straně stojí servery, které poskytují informace a zajišťují služby. Kdysi dávno to byly jen služby spojené s elektronickou poštou a zajištěním přístupu k výpočetním serverům, později se služby rozšířily do sféry bankovní, knihovní, komerční a administrativní. Před několika lety to byl web, který zajistil, že se Internet stal přitažlivým pro širokou uživatelskou komunitu. Pořád to však byl provoz typu klient-server. V posledních letech však i toto schéma bylo překonáno. V podstatě „přes noc“ se začal masivně prosazovat provoz typu P2P („peer to peer“), neboli provoz mezi samotnými účastníky Internetu, bez nutnosti služeb poskytovaných nějakými „centrálními“ servery. Podle některých odhadů tvoří dnes takovéto peer to peer aplikace

50-70 % celkového objemu přenosu dat v Internetu.

Tuto novou éru odstartoval na přelomu tisíciletí program Napster, který umožnil masové přenosy hudebních nahrávek. Jeho popularita během několika málo měsíců vzrostla natolik, že vážně ohrozil celý multimediální průmysl a hudební nakladatelé nakonec dosáhli toho, že Napster musel skončit svoji činnost. To však vůbec neznamenalo konec aplikací typu P2P, spíše naopak. Vývojáři aplikací i uživatelé si uvědomili, jaký vnitřní potenciál Internet má, a že se vůbec nemusí data soustřeďovat na jednom místě, aby byla přístupná mnoha uživatelům. Pokud jsou data zajímavá, pak si je ti, kteří o ně mají zájem, najdou sami - a sami jsou ochotni si je také vzájemně vyměňovat. Legální aspekty výměny dat, které jsou předmětem licencí, je jiná otázka, o té zde nebudeme hovořit. Krátce po „pádu“ Napsteru se objevily jiné programy, které uživatelům zajistily totéž, co dělal Napster. Pracovaly na mnohem promyšlenější strategii, která dovoľovala vytvářet distribuované virtuální společenství uživatelů. Programy DirectConnect, Kazaa, Gnutella, BitTorrent a mnoho dalších jejich odvozenin umožňují přenos dat v dynamicky vytvářených virtuálních sítích, podle příslušnosti ke komunitě dané protokolem přenosu. Většinou jde o přenosy velkých objemů dat, kterými jsou hudební a video nahrávky (MP3, DVD), ale i další objemově náročné informace - databáze, velké softwarové balíky, atd. V mezích pokročila i technologie digitálního zpracování zvuku a videa a jejich přenos po Internetu, a tak se technologie P2P začíná uplatňovat i v této oblasti. Vznikla např. první celosvětová síť Skype, která umožňuje telefonování po Internetu. Skype má dnes kolem 5 milionů uživatelů, a po jeho úspěchu se objevují další sítě podobného typu. Uvedené aplikace P2P jsou samozřejmě ty nejznámější, protože jsou hodně medializovány. Existují však také aplikace, které byly vyvíjeny pro oblasti spojené s metropolitními informačními systémy, výzkumem klimatu apod. Dá se říci, že P2P sítě jsou stejnou revolucí v Internetu, jakou do něho před několika lety přinesl WEB. Jsou to velice spolehlivé distribuované sítě s automatickými funkcemi vyhledání partnera, buď úplně

distribuované nebo jen s minimálními vazbami na nějaké centrum.

Každá z těchto sítí používá trochu jinou strategii vyhledávání partnerů i přenosu dat mezi nimi. To mělo samozřejmě svoji odezvu v akademické oblasti, protože zmíněné problémy vedou k velice zajímavým vědeckým teoriím a řešením obecných problémů, které jsou spojeny s teorií grafů, pravděpodobností, spolehlivostí, plánovacími strategiemi apod. Svědčí o tom velké množství prací, které byly v posledních 5 letech publikovány ve významných časopisech a na odborných konferencích z oblasti počítačových věd. Vytvořit síťovou aplikaci není vůbec jednoduché a pokud by to měla být aplikace, která má fungovat v globálním měřítku, sloužit milionům uživatelů, tak je nutné její funkčnost ověřit v reálném prostředí. Jak ukazují zkušenosti, tak právě „reakce reálného prostředí“ Internetu je katalyzátorem pro ty nejatraktivnější komerční projekty. Vývoj je tak rychlý, že ani velké nadnárodní společnosti a tvůrci komerčních internetových aplikací nemají mnoho času si dělat vlastní základní výzkum v této oblasti a stále více se opírají o teorie, principy, výsledky měření nebo pilotní návrhy, které byly vytvořeny v akademickém prostředí. Dá se říci, že do jisté míry k úspěšnému rozvoji této oblasti přispěla i Planetlab. Existuje dlouhý seznam vědeckých prací, které byly publikovány na základě výsledků z experimentování na PlanetLab. Ve většině těchto prací i v popisu nových projektů dominují problémy související s hledáním informací v rozsáhlých sítích, sdílením a replikací dat, návrhy vhodných struktur a adresací objektů. Často se zde řeší problémy nových metod přenosu dat, využití možných redundantních spojení, což je často (přímo i nepřímo) spojeno s důležitou problematikou směrování.

PlanetLab řeší aplikace budoucnosti

Vyjmenujme si jen několik konkrétních příkladů služeb a aplikací, které v PlanetLab dnes pracují, a které spoluvytváří nové prostředí a tím pomáhají řešit síťové problémy s jistou abstrakcí. Následující příklady jen ilustrují tematiku a ukazují dimenze veličin, se kterými projekty kalkulují:

- Jeden z největších projektů je OceanStore [11]
 - globální paměť, v níž se uživatelé nestarají o to, kde jsou jejich data uložena.
- Distribuční systém Coral <http://www.coralcdn.org> je určen pro efektivní distribuci dat uživatelů, kteří mají pomalé připojení k síti a potřebují informace vystavené na jejich počítačích (webových serverech) distribuovat širokému okruhu dalších.
- Služba CoDNS zajišťuje v síti distribuovaný DNS, který je odolnější proti útokům a nespolehlivosti serverů <http://codeen.cs.princeton.edu/>.
- CoBlitz, CoWeb a CoDeploy jsou služby, které využívají distribuovanou síť pro paralelní přenos velkých objemů dat, např. pro distribuci zdrojů operačních systémů.
- Projekt RON (Resilient Overlay Network) se zaměřuje na tvorbu a optimalizaci směrovacích tabulek s využitím redundance a interního monitoringu dostupných cest v síti [6].
- Projekt DHARMA (Distributed Home Agent for Remote Mobile Access) řeší situace zajištění mobilního přechodu z různých sítí bez nutnosti opětovně navazovat spojení <http://dharma.cis.upenn.edu/>.

Mnoho z těchto služeb a aplikací je přímo svázáno nebo navazuje na další výzkumné projekty, které byly a jsou podporovány nadací NSF (National Science Foundation), jako jsou např. Tapestry, Chord, Pastry, OpenDHT – což jsou projekty patřící do třídy DHT (Distributed Hash Table) [7, 8, 9]. Tyto metody jsou dnes jednou z nejvíce se rozvíjejících oblastí základního výzkumu, a také oblastí s největším publikačním aktivitou. HT – neboli asociativní paměti – zná každý programátor, který pracuje s Perlem a ví, jak užitečné jsou tyto funkce při práci s datovými strukturami. Zjednodušeně se dá říci, že nahrazují dlouhé hledání objektu tím, že objekt má přidělen jednoznačný klíč, který se dá spočítat. Nalezení objektu je pak vlastně ve většině případů realizováno v jednom kroku. Stejný princip se používá i při lokalizaci objektu v síti.

Řada projektů řešených v PlanetLab je podporována nebo realizována s účastí firem jako jsou HP, Microsoft, Intel apod. Například s podporou firmy Intel bylo testováno prostředí pro projekt

IrisNet (Internet-scale Resource Intensive Sensor Network Services), což je síť osobních WEB kamer, u níž se předpokládalo, že by v ní mohlo být zapojeno až 300 milionů lidí! Skromnější modely uvažovaly jen o metropolitních dimenzích s 1,5 milionem lidí, avšak s mnoha typy serverů, které by mohly hlásit stav obsazenosti parkovišť, stav provozu na křižovatkách apod. V tom případě by jen udržování informace o lokalizaci objektů vyžadovalo asi 25000 aktualizací za sekundu [12]. Pro Wikipedia, což je dnes nejrozšířenější světová encyklopedie, se zde testují možnosti efektivní distribuce velkých objemů dat k co nejširšímu okruhu uživatelů s použitím několikastupňové replikace.

Síť PlanetLab používá mnoho uživatelů i pro měření charakteristik chování Internetu – například Google zde má otevřený projekt, který vyhodnocuje dostupnost jejich serverů z různých částí světa. Jiný projekt z této oblasti, který se zaměřil na vzájemnou dostupnost účastníků z různých domén, řeší HP.

Proč je PlanetLab právě tím vhodným prostředím, kde se takovýto výzkum dá provádět? Její přednost spočívá především v tom, že umožňuje uživatelům vytvářet nezávislé aplikace, které mohou běžet v celé síti vedle sebe a vytvářet tak ucelené virtuální vrstvy sítě. Vrstvy používají společné uzly bez toho, že by se při jejich užití jakkoliv vzájemně ovlivňovaly. Tvůrce nebo uživatel aplikace si může v síti vytvořit svoji vlastní strukturu, se kterou pak může pracovat. Nezřídka takové struktury obsahují stovky uzlů PlanetLabu. Pro identifikaci objektů a uzlů v síti používají vlastní adresaci i vlastní metody vyhledávání dat a směrování toku informací mezi objekty a uzly. Některé z nich používají i vícevrstvou architekturu. Použijí některý z dostupných nástrojů typu DHT jako obecný základ pro adresaci a vyhledávání objektů a teprve nad ní budují vlastní aplikaci zpracování dat. Příkladem takového použití je projekt OceanStore nebo projekt Shuttle (bezpečná decentralizovaná distribuce zpráv), který používá v nižší vrstvě systém Tapestry.

Co říci závěrem?

V této první informaci o Planetlabu jsme přinesli několik nových informací o tom, jakým směrem se ubírá výzkum v oblasti síťových technologií ve světě. Je tak trochu překvapivé, že o popisované směry výzkumu se v ČR dosud skoro nikdo nezajímal. Asi jsme příliš malá země, kterou v poslední době zajímala především snaha o to, jak se nejlépe prosadit na gigabitových spojích, a samozřejmě také celosvětová gridomanie.

Síť PlanetLab je zajímavá i z mnoha jiných aspektů. Jedním z nich je třeba to, jak byla taková síť postavena, na jakých principech realizuje svoji virtualitu, a proč se stala modelem pro NGI (Next Generation of Internet). Technické řešení Planetlabu si ukážeme v některém z pokračování tohoto článku.

Neméně zajímavá je také skutečnost, že PlanetLab je velice vstřícná ke všem partnerům, kteří ji chtějí používat. Jejím členem se může stát každá akademická organizace, která do ní vloží svůj uzel. Této příležitosti využil i CESNET; nakoupil požadovanou techniku a podal přihlášku, která se v současné době vyřizuje. Doufáme, že během léta se stihnou realizovat všechny instalační práce a od září bude mít i ČR zastoupení v této unikátní celosvětové laboratoři. V naší akademické komunitě se tím významně rozšíří podmínky pro výzkum v dané oblasti. Není to ale jen záležitost CESNETu, který se přímo zapojí do dnes prováděných experimentů. Je to i výzva pro výzkumné a vědecké týmy z našich vysokých škol a vědeckých pracovišť, aby se začlenily s vlastními projekty. PlanetLab nabízí ideální prostředí pro řešení diplomových nebo disertačních prací i pro navázání přímých kontaktů s předními světovými odborníky.

Literatura

- [1] Andy Bavier, Mic Bowman, Brent Chun, David Culler, Scott Karlin, Steve Muir, Larry Peterson, Timothy Roscoe, Tammo Spalink, Mike Wawrzoniak. Operating System Support for Planetary-Scale Network Services. <http://www.planet-lab.org/>
- [2] P.Barham, B.Dragovic, K.Fraser, S.Hand, T.Harris, A.Ho, R.Neugenbauer, I.Pratt and A.Warfield. Xen and the Art of Virtualization. In Proc. 19th SOSP, Lake Georgie, NY, Oct 2003
- [3] A.Bavier, T.Voigt, L.Peterson, M.Wawrzoniak. SILK: Scout Path in the Linux, Technical Report. 2002-009, Department of Information Technology, Uppsala University, Uppsala, Sweden 2002
- [4] Linux VServers Project, <http://linux-vserver.org/>
- [5] Hary Balakrisnan, Frans Kaashoek, David Karger, Robert Morfia and Ion Stoica. Looking up Data in P2P System. Communications of ACM, February 2003, Vol 46, No. 3
- [6] D. Anderson, H. Balakrisnan, F. Kaashoek and R. Morris. Resilient Overlay Network. In Proc. 18th SOSP, pages 131-145, Banff, Alberta, Canada, Oct. 2001
- [7] I. Stoica, R.Morris, D. Karger, F. Kaashoek and H. Balakrishnan: „Chord: A scalable Peer-to-peer Lookup Service for Internet Applications“. ACM SIGCOMM, San Diego CA, 2001
- [8] V. Ramasubramanian and E.G.Sirer. Behive: Exploiting Power Law Query Distributions for O(1) Lookup Performance in Peer to Peer Overlays. MIT, Cambridge, MA
- [9] Sylvia Ratnasamy, Paul Francis, Mark Handley, Richard Karp and Scotty Shenker. A scalable content-addressable network. In Proc. ACM SIGCOMM 2001, August 2001
- [10] M. Harren, J.M. Hellerstein, R. Huebsh, D.T. Loo, S. Shenker and I. Stoica. Complex Queries in DHT-based Peer-to-Peer Networks. In Proc. 1st Int. Workshop on Peer-to-Peer Systems (IPTPS'02), Cambridge, MA, Mar 2002
- [11] J. Kubiawitz, D. Bindel, Y. Chen, S. Czerwinski, P. Eaton, D. Goels, R. Gummadi, S. Rhea, H. Weatherspoon, W. Wemer, C. Wells and B. Zhao. OceanStore: An Architecture for Global-Scale Persistent Storage. In Proc. 9th ASPLOS, pages 190-201, Cambridge, MA, Nov 2000
- [12] Philip B. Gibbons, Bred Karp, Yan Ke, Suman Nath and Srinivasan Seshan. IrisNet: An Architecture for WorldSensorWeb. IEEE CS and ComSoc. Pervasive Computing, OCT-

DEC 2003

- [13] J. Navrátil, M. Neuman, K. Smejkal. Sít' ČVUT
v roce 2001. Pražská technika, 2001/4, str.
31-36, ISSN 1213-5348. □