

ÚVT MU zpravodaj

Bulletin pro zájemce o výpočetní techniku na Masarykově univerzitě • prosinec 2008 • roč. XIX • č. 2

Dotazník ke Zpravodaji ÚVT MU – vyhodnocení

Miroslav Bartošek, ÚVT MU

Na dotazník vyhlášený v minulém čísle Zpravodaje ÚVT MU došlo během měsíce a půl celkem 118 odpovědí, což hodnotíme jako velmi příznivý výsledek. Vybrané souhrnné výsledky z dotazníkového šetření uvádíme níže. Podrobnější údaje včetně grafů jsou k dispozici v elektronické verzi tohoto článku na webu zpravodaje.

1 Údaje o respondentech

72% z celkového počtu 118 odpovědí poskytli zaměstnanci MU (převážně odborní pracovníci a učitelé); zbytek pochází od studentů bakalářského či magisterského studia (16%) a studentů doktorského studia MU (11%). Jedna odpověď byla od mimouniverzitního čtenáře.

Více než dvě třetiny všech respondentů uvedlo jako své primární profesní zaměření oblast informačních a komunikačních technologií (ICT), což koresponduje i s fakultním rozdělením, kde podle očekávání dominovaly odpovědi z fakulty informatiky (34%) a ÚVT (32%). Mimo tato dvě inforatická pracoviště byly významněji zastoupeny ještě fakulty přírodovědecká (13 odpovědí, tj. 11%), filozofická (6 odpovědí) a lékařská (5 odpovědí).

Z hlediska věkové struktury byli nejvíce zastoupeni mladí uživatelé ve věku do 30 let, kteří tvořili 45%. Ve věku 30-50 let bylo 36% respondentů, věková kategorie nad 50 let byla zastoupena zhruba z jedné pětiny (19%).

2 Vyjádření ke Zpravodaji ÚVT MU

Téměř dvě třetiny respondentů sledují zpravodaj dlouhodobě nebo alespoň několik let, přičemž zastoupení čtenářů tištěné verze a webové verze je vyrovnané: jedna čtvrtina sleduje jen tištěnou verzi a druhá čtvrtina sleduje zase jen webovou verzi (polovina čtenářů sleduje v nějakém poměru obě verze Zpravodaje).

Z hlediska přínosu hodnotí články ve zpravodaji 43% všech odpovědí jako většinou přínosné, a dalších 53% jako občas zajímavé. Jenom 4% odpovědí označilo články za málokdy zajímavé. Forma podání odborných článků vyhovuje naprosté většině respondentů – 82% ji označilo jako „tak akorát“. Zhruba 11% čtenářů shledává odborné články jako příliš technické či málo srozumitelné. Za příliš populární je považují jen 4% respondentů.

Nejvíce zajímají čtenáře následující kategorie článků:

- informace o zajímavých projektech,
- aplikace ICT v různých oblastech,
- informační systémy na MU,
- počítačová bezpečnost,

- popis užitečných softwarových programů a nástrojů.

Tato témata zajímala více než 50 % všech respondentů, kteří na danou otázku odpověděli. Na otázku „které typy článků mne zajímají nejméně“ odpovědělo jen 21 respondentů (18 %), z toho většina označila za nejméně zajímavé úzce zaměřené a příliš technicky pojaté články.

Na otázku „Jaká témata ve Zpravodaji postrádám“ odpovědělo 17 čtenářů, přičemž odpovědi byly velmi různorodé: od popisu software v počítačových učebnách a multilicence MU, přes praktičtější informace o možnostech využití VT ne-specialisty, informace o aplikacích v operačním systému Unix, upozornění na nově připravované technologie a aplikace na MU, popisy mobilních zařízení, až po „počítalogické“ vtípy.

Obecný komentář ke Zpravodaji poskytlo 19 respondentů. Třináct komentářů bylo čistě pochvalných, jeden byl kritický („myslím, že před několika lety býval Zpravodaj zajímavější“). Mezi doporučeními se objevil požadavek na více kratších článků v jednom vydání a větší samostatnost webové verze Zpravodaje (rozšíření co do obsahu i funkce - např. diskuse k článkům).

Zajímavé byly i odpovědi ohledně předplatného tištěné verze: všichni ti, kteří uvedli, že tištěnou verzi odebírají, chtějí v odběru i nadále pokračovat. Naopak, ze 47 čtenářů, kteří tištěnou verzi neodebírají, vyjádřili zájem o její zasílání jen čtyři.

A poslední údaj: získali jsme jednoho nového zájemce o přispívání do Zpravodaje.

Jménem redakce chci poděkovat všem čtenářům, kteří si udělali čas a na náš dotazník odpověděli. Vaše odpovědi byly pro nás užitečné a v mnohém inspirativní. Mohu slíbit, že se vašimi názory a náměty budeme opravdu seriózně zabývat a využijeme je při přípravě dalších čísel Zpravodaje ÚVT MU. □

Služby ÚVT pro vědu a výzkum

David Antoš, ÚVT MU

Ústav výpočetní techniky MU připravuje rozsáhlé rozšíření portfolia služeb, které nabízí uživate-

lům na univerzitě. K zaběhnutým službám, jako jsou například budování a správa počítačové sítě, elektronické informační zdroje, počítačové studovny nebo systémy pro personalistiku a ekonomiku, budou postupně přibývat další, zaměřené zejména na vědecké a výzkumné pracovníky. Hlavními nabízenými službami budou ukládání dat, hosting webových serverů a prostředí pro náročné výpočty.

V tomto článku se budeme věnovat převážně službám ukládání dat - zejména proto, že jsou už dnes dostupné v pilotním provozu a v současné době přecházíme do provozu produkčního.

1 Služby ukládání dat

Proč vlastně hodláme poskytovat ukládání dat jako službu? Nestačí snad koupit si větší disk? Jedním z hlavních úkolů tohoto článku je přesvědčit laskavého čtenáře, že tomu tak docela není.

Uživatelé se v oblasti ukládání dat setkávají typicky s alespoň jedním z následujících tří problémů:

1. velkým objemem dat,
2. požadavky na spolehlivost uložení,
3. potřebou data sdílet mezi skupinou uživatelů, případně je kontrolovaně zveřejňovat.

1.1 Objem dat

Objem dat je měřítkem značně proměnlivým. Do jisté míry lze nedostatek kapacity řešit pořízením většího nebo dalšího disku, v dnešní době není problém pořídit disky kapacity přesahující 1 TB. Pokud ani to nestačí, je třeba se přesunout do kategorie diskových polí, která jsou už nezanedbatelnou položkou rozpočtu, je třeba k nim obvykle pořídit ještě řídicí počítač, nejsou už zrovna nejskladnější a produkují teplo a hluk. Bez příslušného prostředí (což se týká zejména chlazení) se také diskovému poli zkracuje životnost. Celkově lze říci, že diskové pole není právě nejpříjemnější spolubydlící v kanceláři, takové zařízení patří do klimatizovaného počítačového sálu. Také správa diskového pole už není záležitostí pro byt' i znalejšího uživatele, řešení problémů už vyžaduje znalost technologie.

1.2 Spolehlivost uložení

Zatímco klasické knihovny jsou budovány na ukládání na stovky let, nápisy tesané v kameni jsou čitelné i po tisíciletí, moderní technologie zpracování a ukládání dat mohou poskytnout srovnatelnou trvanlivost jen za cenu adekvátní péče o uložené materiály. Trvanlivost uložených dat je daleko skrytějším problémem než záležitost kapacity.

Presvědčení, že uložená data jsou trvalá, vydrží uživateli do havárie disku nebo do prvního CD-R média, které nejde přečíst. Počítačové technologie sice umožňují ukládání velkých objemů dat, jejich ztráta je však většinou o to bolestnější, zejména jedná-li se o data související s měsíci či roky výzkumu. Navíc mírně poškozenou knihu lze bez problémů číst, okopírovat nebo přinejhorším přepsat, z pevného disku spadlého z metrové výšky se nemusí podařit zachránit nic.

Ani jediné diskové pole (třebaže je obvykle odolné například vůči výpadku jednoho z disků), není-li dále zálohováno, ještě nechrání před nechtěným smazáním důležitého souboru nebo třeba před živelnou katastrofou. Rozumnou zálohovací strategií je vytváření dostatečného množství kopií dat. Je zcela zásadní, aby zálohování bylo automatizované, jinak na něj uživatelé budou zapomínat. Toho lze nejnázne dosáhnout, pokud má zálohování formu služby, o kterou se uživatel vůbec nemusí starat, pouze ví, kam se obrátit v případě, kdy o nějaká data přišel a potřebuje je obnovit.

1.3 Sdílení

Posledním zmíněným aspektem je potřeba sdílení. To může mít nejrůznější formu. Často postačí prostý přístup do sdíleného souborového systému pro skupinu uživatelů. Další možností je zpřístupnit část uložených dat definované skupině uživatelů přes webové rozhraní. Tímto způsobem jsou například realizovány dermatopatologické atlasy <http://atlases.muni.cz>. Autor a správce mají přímý přístup do souborového systému, registrovaní uživatelé obrázky a popisy prohlížejí přes webové rozhraní. Je možné zapojit rozličné autentizační mechanismy, pro přístup do atlasů se například

nemusejí zvlášť registrovat uživatelé federací z Dánska a Švédska.

2 Současné možnosti ukládání dat na ÚVT

V současné době nabízí ÚVT datová úložiště jako službu. To zbavuje uživatele závislosti na konkrétní technologii a starosti o ni. Data jsou ukládána na disková pole, což snižuje riziko výpadku živé kopie dat, a jsou zálohována na pásky. Páskové zálohy se provádějí denně a kromě obnovení dat v případě havárie pole umožňují získat omylem smazaná data nebo dohledat starší verzi souboru. Pásky lze po dohodě použít i k dlouhodobé archivaci.

Z pohledu uživatele je úložiště pouze „dalším vysokokapacitním diskem“ v operačním systému, takže není třeba, aby se uživatel učil novou technologii: práci se souborovým systémem každý dobře zná. Standardně poskytujeme připojení protokoly Samba a NFSv4, v odůvodněných případech lze dohodnout i jiné. Úložiště je jako souborový systém přístupné po síti z více míst, například včetně domácích kanceláří. Protože síťové prostředí ne vždy dovoluje přenášet data síťových souborových systémů (je to obvyklé na konferencích, letištích a podobně, kde občas nelze rozumně použít ani služby VPN¹), k datům lze přistupovat i pomocí autentizovaného webového rozhraní. Zde vycházíme z předpokladu, že webový provoz je obvykle na sítích povolen, což dává prakticky neomezenou dostupnost uložených dat.

Úložný prostor vytváříme na žádost podle individuálních požadavků. Lze vytvářet i sdílené diskové prostory vhodné pro menší spolupracující týmy (typicky jednotky uživatelů). Limity objemů dat jsou dohodnuty podle potřeb uživatele, struktura, charakter, užití a publikování uložených dat je na rozhodnutí uživatele, přičemž ÚVT poskytuje konzultační činnost.

3 Webové služby

Mnoho pracovišť na univerzitě provozuje webovou stránku jako prostředí podporující vědeckou spolupráci. Kromě základních informací

¹Pro podrobnější popis univerzitní VPN viz <https://vpn.muni.cz/>, případně článek [2]

o laboratořích zveřejňují i rozsáhlejší data související přímo s výzkumem.

To obnáší správu fyzického hardwaru (který je potřeba mít někde fyzicky umístěn), instalace operačního systému, webového serveru, databáze, systému pro správu obsahu (Content Management System, CMS) a konečně vlastního obsahu webové prezentace. Přitom pouze obsah prezentace má odborný obsah odpovídající činnosti pracoviště, vše ostatní představuje zátěž.

Plánujeme proto hosting webových serverů, zahrnující všechny součásti od fyzického hardwaru po systém pro správu obsahu. Předpokládáme nasazení několika běžných CMS. Pokud uživatelům nebude tato nejvyšší vrstva systému vyhovovat, mohou použít CMS vlastní, který si budou spravovat sami. Mohou však použít všechny vrstvy až po databázi. Podobně bude možno rozdělit zodpovědnost i mezi nižší dvojici vrstev.

Stránky budou přístupné pod doménovými jmény třetího nebo vyššího řádu (např. <http://sito1a.fi.muni.cz>), podporovat budeme i napojení nezávislých domén (<http://www.sito1a.cz/>). Zásadním prvkem webových služeb je úzké provázání s datovými úložišti, které dovolí uživatelům zveřejňovat uložená data ve formě, kterou považují za nejvhodnější pro konkrétní případ. Přístup k datům lze navázat na různé autorizační a autentizační mechanismy jako jsou federace apod.

Pro efektivní využití hardwaru a logické oddělení jednotlivých webových serverů bude použita technologie virtuálních počítačů. Využití virtuálních počítačů je ovšem daleko obecnější, jak uvidíme v následující části.

4 Prostředí pro náročné výpočty

Skupiny uživatelů, které potřebují větší výpočetní výkon, než poskytuje běžný osobní počítač, často řeší situaci pořízením vlastního clusteru. Zásadními nevýhodami tohoto přístupu jsou zejména nevyužívání dostupných zdrojů v době, kdy není velká potřeba výpočetního výkonu, a na straně druhé nedostatek výpočetních zdrojů ve fázích, kdy je uživatel intenzivně potřebuje, protože potřeba výpočetního výkonu má typicky nárazový charakter. Spotřeba energií,

stejně jako úsilí nutné ke správě clusteru, jsou ovšem téměř nezávislá na míře využití clusteru. Obvyklým a velmi nesnadno řešitelným problémem je také fyzické umístění počítačů: výkonné stroje produkují teplo (a tím nutné návazné investice do chlazení) a hluk.

Na druhou stranu vlastní cluster umožňuje uživatelům provozovat prostředí vyladěné pro jejich potřeby. Pokud ovšem uživatel své výpočty provádí na pronajaté infrastruktuře, při klasickém přístupu s uživatelskými účty na sdíleném operačním systému je omezen na prostředí, které připravil provozovatel takové infrastruktury, na jím poskytovanou distribuci operačního systému a dostupný software.

Virtualizace počítačů (pro úvodní přehled tohoto tématu viz [1] a následující díly seriálu) umožňuje odstranit většinu výše uvedených nevýhod současně, logicky odděluje instalaci operačního systému a softwaru od fyzického stroje a dovoluje spouštět celé instalace operačních systémů.

Uživatel může dokonce mít v takovém prostředí správcovská práva k operačnímu systému, takže se nevzdává možnosti použít vlastní prostředí a plně jej spravovat. Předpokládáme samozřejmě poskytování standardní instalace pro uživatele, kteří nechtějí investovat úsilí do vytváření instalace vlastní.

Výpočetní zdroje budou dostupné na žádost, uživatel si vyžádá určitý počet strojů s danou instalací na určitou dobu a počet zapojených strojů lze během života clusteru měnit. Protože jsou s vysokou pravděpodobností špičkové výpočetní potřeby jednotlivých uživatelů rozloženy v čase, celková fyzická infrastruktura pro virtuální prostředí je menší, než celkové množství strojů potřebné pro jednotlivé uživatele. To přináší celkové zvýšení efektivity a spolehlivosti systému a energetické úspory. Nárazově lze používat i daleko větší výpočetní kapacitu, než by si uživatel mohl dovolit pořídit jako vlastní cluster.

V každém případě se ovšem uživatel zbaví nutnosti spravovat fyzické počítače a základní infrastrukturu (napájení, chlazení, sítě, fyzické zabezpečení serverovny, ...), plánované výpočetní zdroje budou umístěny na sálech ÚVT se vším potřebným vybavením a zajištěnou správou.

V současnosti celý systém virtuálních výpočetních zdrojů analyzujeme a připravujeme technologie.

5 Shrnutí

Zatímco koncepce používání virtuálních clusterů se teprve připravuje, „vedlejší produkty“ tohoto úsilí se ukazují jako užitečné už dnes. Úložné kapacity jsou nutnou součástí takové infrastruktury, webové servery jsou z technologického hlediska předstupněm ke kompletnímu výpočetnímu prostředí. Úložné kapacity byly v roce 2008 k dispozici vytipovaným pracovištím univerzity, se kterými jsme jednali individuálně v rámci analýzy požadavků, a získaly během tohoto období přes 35 uživatelů, kteří je používají k nejrůznějším účelům. Příkladem může být ukládání videa na LF, sdílený repozitář fotografií historických listin na FF, nebo třeba archiv výsledků experimentů z PřF.

Pokud máte zájem zvýšit spolehlivost a dostupnost uložení svých dat, navštivte náš web <http://storage.ics.muni.cz> a kontaktujte nás na e-mailové adrese storage@ics.muni.cz. Na stejné adrese také přivítáme jakékoli požadavky, přání, nápady a komentáře k celému projektu výpočetních a datových zdrojů a dalších virtualizovaných prostředí.

Literatura

- [1] Luděk Matyska. Virtualizace výpočetního prostředí. *Zpravodaj ÚVT*, XVII(2):9–11, 2006. ISSN 1212-0901.
- [2] Jakub Morávek a Radim Peša. VPN server Masarykovy univerzity. *Zpravodaj ÚVT*, XIV(2):10–12, 2003. ISSN 1212-0901. □

Velkoformátové tisky na MU

Kamil Malinka, Jiří Ledvinka, ÚVT MU

Masarykova univerzita je druhou největší veřejnou vysokou školou v České republice a největší na Moravě. Jednu ze základních priorit MU tvoří věda a výzkum. Ten s sebou přináší mimo jiné i potřebu prezentovat výsledky výzkumných projektů. Jednou z častých forem prezentace je

poster – velkorozměrný plakát, prezentující výsledky práce formou tabulek, fotografií a komentovaných témat. Kromě výzkumu lze nalézt mnoho dalších oblastí, jež potřebují připravovat velkoformátové tisky. Jedná se například o výstupy z geografických informačních systémů, pozvánky a prezentace na nejrůznější akce apod.

V současné době je tisk posteru (a velkoformátového tisku obecně) poměrně nákladnou záležitostí, jež se řeší u specializovaných externích firem. Na základě zaznamenaných potřeb se ÚVT rozhodlo nabídnout službu, která zpřístupní uživatelům naší univerzity velkoformátový tisk za výhodných podmínek přímo na ÚVT MU. Plotter bude k dispozici jak široké akademické obci, tak i studentům.

1 Plotter a jeho možnosti

Pro velkoplošné tisky byl zakoupen plotter HP Designjet Z6100ps 42 palců. Je vybaven šesti tiskovými kazetami a umožňuje kvalitní barevný tisk o rozlišení až 2400 × 1200 dpi (optimalizovaných ze vstupních 1200 × 1200 dpi). Jako tisková technologie je použit termální inkoustový tisk HP. Plotter umí tisknout na velké množství typu papírů, které se do tiskárny vkládají po rolích. Maximální šířka média je 1067 mm, s maximální tisknutelnou plochou omezenou na 1057 mm. Délka média může dosahovat až 175 m v závislosti na druhu média. Plotter umí tisknout jakékoli rozměry, a není tedy třeba striktně dodržovat standardizované formáty. Největší síla média může být až 430 g/m² (síla běžného kancelářského papíru je 80–90 g/m²). Kromě tisku na papír rozmanitého druhu je možno tisknout i na speciální materiály jako je např. fólie, textilie či umělecký materiál.

2 Kvalita tisku

Pro dosažení požadované kvality vytištěného projektu je nutný výběr vhodného tiskového média, který je závislý na tiskové úloze. Například není vhodné tisknout fotografii v nejvyšší kvalitě na běžný kancelářský papír, protože papír nepojme tolik inkoustu a zvlíní se. Záleží také na tom, k čemu má vytištěný dokument sloužit. Např. plakát obsahující drobný text a plán budovy lze

pro pracovní potřeby vytisknout na kancelářský papír; k prezentaci na konferenci však již nemusí být takový výstup dostatečně kvalitní. Typ tiskového média (papíru) má na výsledek vytištěného projektu obrovský vliv. To by měl mít každý zákazník na paměti ještě dříve, než zadá konečný požadavek pro tisk. Výběr vhodného média je možné konzultovat s autory článku, kteří rádi poskytnou příslušná doporučení.

3 Tiskové skupiny

Při vlastním tisku zařazuje operátor tiskovou úlohu do vhodné tiskové skupiny provázané se zvoleným médiem. Základní tiskové skupiny jsou následující (seřazené dle náročnosti tisku – od nejméně náročného po nejnáročnější):

Skupina A: plocha s malou hustotou čar. Vhodné např. pro tisk výstupů z geografických informačních systémů (plány budov apod.).

Skupina B: plocha s velkou hustotou čar a popř. i drobné plochy. Např. plány pozemků, které mají větší množství ploch.

Skupina C: mapy, plakáty a interpretace zabírající téměř celou plochu. Např. plakáty, velkoplošné prezentace.

Skupina D: tisk nejnáročnějších úloh na speciální média. Např. fotografie a plakáty tištěné na fotografický papír.

4 Další parametry tiskárny

Rychlost tisků je závislá na druhu tiskové úlohy a pohybuje se v rozmezí 60 m²/h (ekonomický tisk) až 8,1 m²/h (tisk fotografií). Tiskárna je vybavena 1Gbit ethernetovou přípojkou, což umožňuje rychlé nahrávání i rozsáhlých tiskových úloh. Ovládání tiskárny probíhá přes webové rozhraní. Mezi další vybavení tiskárny patří 40GB interní disk, na který jsou ukládány tiskové úlohy.

Tiskárna mimo jiné poskytuje mechanismus pro účtování tisků. Eviduje přesné údaje o spotřebě materiálu pro každou jednotlivou úlohu. Tyto údaje jsou pak použity pro výpočet přesné výsledné ceny tisku.

5 Problematika barevných profilů

Každý uživatel jistě chce, aby mu tiskárna vytiskla projekt přesně tak, jak jej vidí na obrazovce. Ať už se jedná o barvy či průpravu. Zdá se to být triviální problém, ovšem opak je pravdou. Data pro tisk jsou nějakým způsobem uložena v počítači. Počítač tato data nějak zobrazuje a tiskárna tato data nějak tiskne. Jsou zde tedy tři aspekty, které musí být dokonale sladěny, aby se dosáhlo požadovaného výsledku. K tomu slouží barevné profily, které jsou schopny tyto tři aspekty zkorigovat. K plotteru bylo zakoupeno zařízení na kalibraci barev, a tím se z velké části vyřešilo správné zobrazování i tisk barev. Protože se jedná o problematiku, se kterou se zatím seznamujeme, jsme schopni zaručit přesnou korelaci barev při tisku a počítačovým zobrazením zatím pouze na našem pracovišti s kalibrovaným monitorem. Pokud zákazníkovi záleží na přesném vytisknutí barev, je zatím nutné si barvy před tiskem zkontrolovat na našem zkalibrovaném pracovišti.

6 Jak službu využívat

V současnosti byl již dokončen testovací provoz v rámci ÚVT a službu otevíráme pro univerzitní veřejnost. Zájemci z MU o službu velkoformátového tisku necht' laskavě kontaktují autory tohoto článku na adrese plotter@ics.muni.cz, kteří jsou schopni jak poskytnout potřebné konzultace, tak provést kompletní tiskovou zakázku.

K dispozici je standardně 10 druhů papíru různých hmotností a délek. Podle potřeb lze papíry doobjednat a nabídku rozšířit. Kvůli dlouhým dodacím lhůtám je nutné hlásit speciální požadavky co nejdříve. Tisk jsme schopni zajistit z předloh libovolných počítačových formátů. Pro nejlepší výsledky ovšem doporučujeme využít standardních formátů PDF nebo JPG. Pro finální úpravy velikosti tisků je připravena velkoformátová řezačka o maximální řezné ploše 1300 mm. Pro bezpečný přenos vytištěných úloh jsou k zapůjčení tubusy nejrůznějších velikostí.

Jak to vypadá s cenou: Cenu tisku jsme schopni odhadnout s rozumnou přesností ještě před

vlastním tiskem (je nutno dodat, že cena černobílého tisku je stejná jako barevného). Po vytištění úlohy spočítá systém přesnou kalkulaci. Platba probíhá bezhotovostním způsobem v rámci univerzitního systému SUPO (systém úhrady pohledávek za osobami). Zákazník obdrží potvrzení o platbě použitelné pro zúčtování platby v rámci jeho účetní zakázky, například zakázky pro výzkumný projekt.

Příklad cenové relace tisku: Externí firma si v současnosti účtuje za plakát o rozměrech 800 mm × 1000 mm typicky kolem 720 Kč bez DPH. Za stejný tisk na ÚVT zaplatíte zhruba 380 Kč.

Cílem projektu je nabídnout levnější alternativu velkoplošných tisků pro studenty a zaměstnance MU, pro jejich studijní a výzkumné potřeby. Slibujeme si od něj zlepšení možností prezentace výsledků vědeckých skupin i nejrůznějších oddělení MU, stejně jako rozšíření nabídky služeb pro studenty. Příkladem studentského uživatele služby může být grafik, kterému odpadne velká starost se zajištěním tisku jeho práce.

Dosavadní zkušenosti ukazují, že plotter na ÚVT šetří čas a peníze akademické obci MU. Ukázky posterů vytištěné na našem plotteru můžete vidět na chodbách ÚVT MU. □

Osm let Celouniverzitní počítačové studovny

Marcela Valentová, ÚVT MU

1. září 2000 byla v prostorách Lékařské fakulty MU na Komenského náměstí otevřena Celouniverzitní počítačová studovna. Její vznik inicioval tehdejší rektor Masarykovy univerzity prof. RNDr. Jiří Zlatuška CSc. – studovna měla výrazně zlepšit možnosti přístupu studentů k informačním zdrojům na internetu (včetně tehdy nově budovaného IS MU) i vlastní práce s počítačem obecně. Omezený počet počítačů měl být kompenzován nepřetržitým přístupem k nim, což byla ve své době v českém akademickém prostředí rarita. Rektorova iniciativa se setkala s rozpačitostí velké části akademické veřejnosti, neboť s provozem takového zařízení nebyly žádné zkušenosti. Bylo slyšet optimistické, ale i pesimistické výroky typu “studenti raději půjdou

do hospody než do studovny“. Přesto však byla studovna začleněna do probíhající rekonstrukce uvedené budovy. Navržením, realizací a provozem byl pověřen Ústav výpočetní techniky MU (ÚVT).

Výsledkem intenzivního úsilí bylo vybudování pracoviště se 109 PC s procesorem Intel Celeron 533A, 128 MB RAM, HDD 10GB, 17“ monitory, s disketovou a CD-ROM mechanikou. Operačním systémem byl Windows 2000 Professional CZ s aplikacemi.

Zájem studentů o toto pracoviště však předčil veškerá očekávání, dokonce i nepřetržitý provoz, o jehož opodstatnění se zpočátku mnohdy pochybovalo, se osvědčil.

Postupně byly získávány zkušenosti jak s provozem instalované výpočetní techniky i souvisejících technických zařízení, tak se samotným provozem spočívajícím např. ve sledování chování uživatelů. Rovněž byly získávány zkušenosti ve vztahu k uživatelům. Bylo nutné řešit nejrůznější prohrěšky proti provoznímu řádu a postupně je eliminovat na minimum pomocí technických i organizačních prostředků. Zájem studentů ovlivňovat provoz studovny vyústil ve zřízení Spolku CPS, který vytvářel vedení ÚVT a MU zpětnou vazbu o účelnosti tohoto zařízení. Na základě toho všeho bylo pracoviště dovybaveno klimati-začními jednotkami, zdokonaleným kamerovým a zabezpečovacím systémem a systémem monitorování stanic.

V prvních 3 letech provozu byla roční návštěvnost 300 000 vstupů uživatelů. Tomu bylo nutné podřídit i inovaci výpočetní techniky, a proto byla v roce 2003 provedena kompletní výměna PC za nové s 512 MB, 2,2 GHz., 19" CRT monitory a 17" LCD panely. Čtvrtina počítačů byla vybavena kombinovanou mechanikou DVD/CDRW. V souvislosti s tím bylo možné rozšířit nabídku programového vybavení, která potom obsahovala dnes již běžné standardy jako Microsoft Office, Statistica, Lingea Lexicon, Matlab a další. Pořízen byl další uživatelský server disponující takovou diskovou kapacitou, aby uspokojil rostoucí nároky uživatelů. Individuální diskový prostor, který je studentům k dispozici, tak mohl být navýšen na 100 MB z původních 15 MB. Postupem času, především na základě požadavků a

zájmu studentů, vyplynula potřeba poskytování dalších doplňkových služeb, souvisejících s provozem CPS. Jednalo se především o tisky, kopírování a drobný prodej záložních médií. Postupně se nabídka rozšířila o materiál na kroužkovou vazbu, kterou si zde mohou studenti sami svázat, dále o termovazbu, laminování, CD se SW Statistica atd. Tato činnost však s sebou přinesla vznik hotovostních plateb, pro které na pracovišti nebyly potřebné podmínky. Proto byl vyvíjen a následně na CPS zaveden bezhotovostní prodej těchto služeb pro potřeby studia. V současné době je tisk a drobný prodej na CPS zařazen do systému SUPO. Tento systém umožnil propojení plateb za drobný prodej či tisk na CPS a ubytování na SKM. V současné době je již 7 bankovníků, které jsou připojeny na systém SUPO (CPS, KUK, SKM – 3x, PedF – 2x).

Neustálý nárůst počtu uživatelů CPS, který měl za následek např. i mnohametrové fronty před vstupem do studovny, přiměl vedení MU ke zřízení dalších počítačových studoven na jednotlivých fakultách. Koncept byl opět pověřen ÚVT a podmínky provozu těchto pracovišť byly zakotveny do směrnice rektora č.13/05. V rámci systému *Univerzitních počítačových studoven* (UPS) tak postupně vznikaly: v Knihovně univerzitního kampusu (KUK), v knihovně na FSS, Informační centrum Přírodovědecké fakulty, studovna v Ústřední knihovně FF a na Pedagogické fakultě. Výsadní postavení mezi nimi však i nadále zaujímá CPS, která i díky své výhodné poloze ve středu města Brna jako jediná poskytuje 24 hodin denně 7 dní v týdnu komfortní služby všem studentům a zaměstnancům MU, a to nejen prostřednictvím výpočetní techniky a tisku, ale i technické podpory. To všechno byly důvody, které opět vedly k nutnosti přemýšlet o zvýšení kapacity CPS. Určitým řešením bylo zavedení bezdrátové Wifi technologie (VPN, později také Eduroam), která umožňuje připojování uživatelů s vlastními notebooky. Ale podstatného zvýšení kapacity bylo dosaženo až rekonstrukcí CPS v roce 2006, kdy byla podle nového architektonického návrhu vybavena 129 stálými PC pracovišti a 16 pracovišti pro notebooky. Na CPS jsou nyní k dispozici PC s procesorem Intel Pentium D 2.8GHz, 1 GB RAM, HDD 80 GB, disketo-

vou a DVD+-RW mechanikou a 19" LCD monitorem. Během provozu CPS se objevily nejrůznější způsoby neoprávněných vstupů formou zneužití přístupových karet či různých způsobů překonávání vstupního zařízení, proto při této kompletní rekonstrukci CPS došlo i k novému zabezpečení vstupních dveří studovny, k výměně turniketu a doplnění kamerového systému. V návaznosti na budování UPS byl nahrazen původní systém monitorování studentů v CPS systémem novým, společným pro všechny tyto studovny. Současně došlo ke zlepšení zázemí pro studenty – byly pořízeny nové automaty na teplé či studené nápoje a na bagety. Přibyly nové tiskárny a kopírky, pracovní prostor pro vazbu materiálů a skartovačka.

S rozvojem studoven na fakultách i se zlepšováním ekonomické situace studentů, kteří ve stále větší míře využívají pro běžné uživatelské činnosti vlastní notebooky, došlo k částečnému odlehčení provozu CPS. S výhledem do budoucna by proto měla CPS své služby zaměřit především na poskytování speciálního a náročnějšího programového vybavení a na vysokorychlostní připojení k internetu. V neposlední řadě stojí i podpora zapojování dalších studoven do systému UPS. Jednak je centrální správa těchto pracovišť ekonomicky výhodnější, a pro studenty je nespornou výhodou stejné základní programové vybavení ve všech těchto studovnách (při zachování výhod specializovaných aplikací v jednotlivých studovnách), a také je kladem stejné pracovní prostředí a přístup k jejich dokumentům.

O setrvalém zájmu o CPS svědčí i vysoké počty návštěvníků v jednotlivých letech (údaj za rok 2000 představuje návštěvnost jen za třetinu roku, od otevření studovny v září 2000):

2000	110 000
2001	387 600
2002	382 100
2003	396 600
2004	407 200
2005	388 200
2006	348 000
2007	378 819

□

Infrastruktura univerzitních počítačových studoven

Jakub Dobrovolný, Vít Bukač,
FI a ÚVT MU

System univerzitních počítačových studoven (UPS) představuje jednotně spravovanou síť počítačových studoven na různých místech MU, která nabízí uživatelům v každém svém bodě jednotné pracovní prostředí, přístup ke svým datům a standardizovanou sadu služeb [1]. Uživatel může podle svých aktuálních potřeb navštěvovat kteroukoliv ze studoven a využívat veškeré jí nabízené informační technologie a služby. V libovolné učebně fungující v režimu UPS mají všichni přístup ke svým centrálně uloženým datům, se kterými mohou pracovat v jednotné množině aplikací sahající od textových editorů až po speciální matematické nástroje.

Základní technologií používanou pro správu tohoto prostředí je *Microsoft Active Directory*. Doménová struktura se skládá z forest root domény a čtyř child domén, z nichž je jedna vyhrazena právě pro univerzitní počítačové studovny. Root doména obsahuje uživatelské účty všech studentů a zaměstnanců Masarykovy univerzity, zatímco v child doménách jsou soustředěny pouze počítače, ke kterým je možné se pomocí těchto účtů přihlašovat. O správu na nejvyšší úrovni se stará malá skupina doménových administrátorů, zaměstnanců Ústavu výpočetní techniky MU, naproti tomu jednoduché a opakované úlohy (instalace stanic, lokální změny politik) vyřizují lokální správci každé z child domén (zaměstnanci jednotlivých zařazených fakult).

Uživatelé

Centralizace správy podle zásady 1 člověk = 1 účet výrazně usnadňuje správu uživatelských účtů. Přidávání a zakazování účtů je automatizované, provádí se několikrát denně na základě změn v Informačním systému MU. Studenti jsou rozdělováni do skupin podle fakulty, oboru atd. až do úrovně jednotlivých předmětů, zaměstnanci až do úrovně pracovišť.

Častým problémem je existence *hostů* (guests), tedy osob, které z nějakého důvodu potřebují

přístup k informačním službám, ale přitom nemají s organizací žádný bližší vztah (např. návštěvníci knihoven, návštěvníci konferencí). V takovém případě je této osobě díky jednoduché webové aplikaci vytvořen nový účet na pevně danou dobu. Tento účet má práva běžného uživatele a přístup ke všem obvyklým zdrojům (Wi-Fi, VPN,...) s výjimkou tiskáren. Záznam o existenci hosta je uchován i poté, co je účet smazán.

Po přihlášení uživatele ke stanici v libovolné učebně v režimu UPS jsou mu zpřístupněna jeho data a profil. Při odhlášení jsou veškeré změny propagovány zpět na server. Data jsou souběžně uložena na dvou diskových polích Hewlett-Packard o celkové kapacitě 2 TB, přičemž každému uživateli je vyhrazen diskový prostor o velikosti 100 MB. Kvůli vzrůstajícím nárokům je však v plánu nasadit do léta 2009 nové úložiště a kvótu navýšit.

Počítačové stanice univerzitních studoven

V současné době je v systému univerzitních počítačových studoven Masarykovy univerzity provozováno více než 500 počítačů, které se nachází v šesti lokalitách (Celouniverzitní počítačová studovna MU na Komenského náměstí, Fakulta sociálních studií, Filozofická fakulta, Přírodovědecká fakulta, Pedagogická fakulta a knihovna univerzitního kampusu). Aby se předcházelo nutnosti dojíždění doménových administrátorů na uvedené lokality, starají se o ně lokální fakultní správci. Jedná se převážně o instalaci/reinstalaci stanic, hardwarovou podporu, správu software, případně řešení problémů uživatelů.

Servery v doméně UCN

Technologie Microsoft Active Directory je v každé doméně nasazena vždy na dvou doménových řadičích (v případě výpadku jednoho přebírá jeho úlohu druhý). Doménoví administrátoři mají přístup ke všem serverům. Lokální administrátoři mají k dispozici třetí *member server* včetně předinstalované aplikace Adminpack umožňující správu infrastruktury bez nutnosti přístupu přímo na stroje. Tento přístup má značně omezená práva, konkrétně se jedná pouze o tvorbu a správu skupinových politik

a úpravu nebo zakazování a povolování účtů (nikoliv jejich vytváření nebo mazání). Na svém member serveru jsou tito správci lokálními administrátory, tudíž je mohou využívat pro vlastní potřeby, např. spravování software typu antivir.

Instalace OS a aplikací

Instalace operačních systémů probíhá pomocí technologie *Unattended* spolu se síťovými kartami s podporou PXE (Preboot Execution Environment). Při startu stroje administrátor zvolí bootování ze sítě, zadá heslo a stanice si stáhne všechny potřebné soubory a bezobslužně nainstaluje systém. Soubory jsou uloženy na datovém úložišti tří serverů, které tvoří cluster (tím je zajištěna vysoká dostupnost v případě poruchy). Zde se nachází kopie instalačních souborů ve všech konfiguracích potřebných pro počítače používané ve studovnách.

Po instalaci se stroj pomocí skriptu přidá do domény a po potřebném restartu se aplikují příslušné politiky a nainstaluje se vybraný software. Pro instalaci software je využita technologie *Microsoft Installer*. Většina MSI balíčků je připravena enterprise administrátory pomocí administrativní instalace. Někteří výrobci dodávají se svými aplikacemi už hotový balíček, který stačí stáhnout a je připraven k použití.

V root doméně je zřízeno centrální úložiště (Distributed File System - DFS), a to se replikuje do každé child domény. Balíčky jsou tak většinou přístupny po rychlejší síti LAN a nezatěžují tolik celkový provoz na síti, pouze v případě výpadku serveru se MSI balíček stáhne z centrálního úložiště.

Instalace MSI balíčků je centrálně spravována pomocí *Group Policy Management*. Vytvoří se politika, v níž je nadefinována instalace, a lokální administrátoři sami rozhodují, který software pomocí dané politiky chtějí mít k dispozici pro uživatele a povolí její aplikování.

Práce z domova

Studenti mají možnost používat vybrané aplikace (z důležitějších můžeme jmenovat Matlab R2007a, STATISTICA 8, ArcView GIS 3.0, ArcGIS

atd.) i mimo studovny MU. Každý student má možnost připojit se pomocí *Vzdálené plochy* na server `tserver.ucn.muni.cz` [2], kde se přihlásí pomocí svého UČA a sekundárního hesla. Zde má již k dispozici svůj profil a svoje data a může pracovat i z domova tak, jako by seděl u stanice přímo ve studovně.

Bezpečnost

Základním principem bezpečnosti v takovéto infrastruktuře je nedávat uživatelům větší práva, než potřebují ke své práci. Proto mohou všichni uživatelé pracovat pouze s běžnými, nikoliv administrátorskými právy, a dopad případného útoku na systém je tak minimalizován.

Podstatným bodem je správa aktualizací. Již samotné instalační soubory jsou udržovány aktualizované. Všechny stanice jsou do infrastruktury zapojeny již s nejnovějšími záplatami. Následná distribuce záplat je řízena ze dvou univerzitních WSUS serverů. Servery WSUS udržují přehled o již aplikovaných updatech a umožňují specifikovat, které záplaty mají být uplatněny na které počítače. Skript kontroluje pravidelně každý týden stav všech strojů, a pokud některému z nich chybí důležité aktualizace, jsou jeho lokální správci informováni e-mailem.

Centrálně je řízena také ochrana proti virům. Všechny stanice v univerzitních počítačových učebnách mají nainstalován antivirový software NOD32, jehož správa a distribuce virových definic probíhají ze serveru pod správou doménových administrátorů. Posledním ze základních lokálních softwarových bezpečnostních prvků je firewall integrovaný v použitém operačním systému Windows XP, jehož nastavení je prováděno pomocí technologie Group Policy.

Součástí zajištění bezpečnosti UPS jsou i opatření pro bezpečnost fyzickou, která v sobě zahrnují především omezení přístupu do prostor s výpočetní technikou pouze pro osoby s platným ISICem nebo zaměstnaneckou kartou, a to buď přes turniket (případ Celouniverzitní počítačové studovny) nebo pomocí čteček karet. Počítače jsou proti krádežím zajištěné zvukovým alarmem, který se spustí, pokud je počítač odpojen z elektrické sítě.

Literatura

- [1] R. Peša, O. Krajíček, L. Rychnovský. *Počítačové studovny MU*. Zpravodaj ÚVT MU. ISSN 1212-0901, 2005, roč. XVI, č. 1, s. 9-11.
- [2] L. Rychnovský, P. Babinec, P. Tuček. *TServer - terminálový server UCN*. Zpravodaj ÚVT MU. ISSN 1212-0901, 2008, roč. XVIII, č. 5, s. 9-11.
- [3] Univerzitní počítačové studovny - webové stránky. <http://www.muni.cz/ics/services/ups/about> □

Souborný katalog ČR

Hana Vochozková, ÚVT MU

Potřebujete nutně určitou knihu či jiný dokument a vaše „mateřská“ knihovna ho nemá ve fondu nebo je zrovna vypůjčen? Místo časově náročného prohledávání katalogů jednotlivých knihoven můžete zkusit prohledat katalog pouze jeden - *Souborný katalog ČR* <http://skc.nkp.cz>. V případě, že vaše hledání bylo úspěšné, můžete pak požádat o zapůjčení dokumentu prostřednictvím MVS (= Meziknihovní Výpůjční Služby).

V Souborném katalogu ČR najdete záznamy o dokumentech, které vlastní české knihovny (vědecké, vysokoškolské, odborné i veřejné), které do Souborného katalogu ČR zasílají svoje záznamy nebo poskytují informace o odběru periodik. Přehled zúčastněných knihoven lze získat na adrese <http://www.caslin.cz/prouzivatele/spolupracujici-knihovny/>.

V roce 2008 do Souborného katalogu ČR aktivně přispívalo (přispívání je dobrovolné) více než 150 českých knihoven. Knihovny MU se mezi aktivní přispěvatele zařadily letos v říjnu. Protože se záznamy do Souborného katalogu ČR zasílají dávkově po částech a podléhají zevrubné kontrole, předpokládáme převedení celého katalogu MU (přes tři čtvrtě milionu záznamů) do Souborného katalogu ČR v druhé polovině roku 2009.

Předchůdcem Souborného katalogu ČR byly souborné katalogy CEZL (Centrální Evidence Zahraniční Literatury) vytvářené od roku 1965 ve Státní knihovně ČSR (nyní Národní knihovně ČR)

v Praze, v té době pochopitelně v klasické tj. listkové podobě.

Zásadní automatizace systému CEZL byla zahájena v rámci projektu CASLIN v první polovině 90. let minulého století. V elektronické podobě je katalog budován od roku 1995, od roku 1997 přijímá i záznamy domácí produkce. Od počátku roku 1998 je Souborný katalog ČR zdrojem pro kooperativní zpracování České národní bibliografie.

V listopadu 2008 obsahoval Souborný katalog ČR cca 3,4 milionu záznamů dokumentů (monografií, speciálních dokumentů tj. hudebnin, map, zvukových dokumentů a seriálů) a to českých i zahraničních bez omezení vročení. Každému záznamu odpovídá obvykle řada knihovních jednotek v knihovnách zapojených do souborného katalogu.

Protože Souborný katalog ČR je stejně jako katalog MU budován v systému Aleph500, nemělo by vyhledávání činit uživatelům MU žádné potíže. V Souborném katalogu ČR (<http://skc.nkp.cz>) lze vyhledávat v celé bázi nebo v dílčí bázi Monografie a speciální dokumenty nebo v dílčí bázi Seriály. Základní vyhledávání umožňuje vyhledat dokument podle názvových údajů, autora, ISBN/ISSN/ISMN, roku vydání, nakladatele. Vyhledávání lze omezit kódem země vydání a jazykem dokumentu. Databázi si lze prohlížet též pomocí rejstříků, v nichž lze vyhledávat podle následujících položek: název, předmětová hesla, autor (příjmení jméno), korporace/akce, nakladatelé, edice, ISBN/ISMN, ISSN, MDT. Je však třeba počítat s tím, že v záznamech Souborného katalogu ČR je tzv. věcný popis velmi různorodý.

Záznamy obsahují informaci (sigla a zkrácený název knihovny) o tom, která ze zúčastněných knihoven dokument vlastní a kde si tedy můžete dokument vypůjčit. Z přehledu vlastníků záznamu lze přejít na záznam příslušné knihovny v bázi ADR (Adresář knihoven a informačních institucí v ČR), kde najdete kromě různých informací o knihovně a kontaktů také funkční link na webové stránky knihovny a na její online katalog.

U řady záznamů je možno propojit se přímo do lokálních katalogů knihoven (vlastníků dokumentů) a získat tak aktuální informaci o mo-

mentální dostupnosti dokumentu v konkrétní knihovně (např. zda je dokument v knihovně k dispozici nebo je právě vypůjčen).

V záznamech seriálů jsou kromě toho uvedeny také údaje o odebíraných ročnících a době uchování. U některých titulů seriálů je možné propojení na plný text.

Přímo ze Souborného katalogu ČR lze odeslat objednávku MVS na výpůjčku dokumentu nebo objednávku na zhotovení kopie dokumentu nebo jeho části. Tato služba je určena pouze knihovnám nikoliv jednotlivcům. Přehled knihoven poskytujících tuto službu naleznete na adrese <http://www.caslin.cz/pro-uzivatele/spolupracujici-knihovny/>.

Za MU je to zatím Ústřední knihovny Pedagogické fakulty, Přírodovědecké fakulty a Fakulty sociálních studií. Zapojení ostatních knihoven MU se předpokládá v nejbližší době.

Literatura

[1] Web souborného katalogu ČR (informační stránky). <http://www.caslin.cz/> □

Nástroje a služby Google. 2. Google Scholar

Miroslav Bartošek, ÚVT MU

Téměř přesně před čtyřmi lety, v listopadu 2004, představila firma Google svůj nový produkt - vyhledávač vědecké literatury *Google Scholar* (GS). Jeho uvedení vzbudilo velkou pozornost a rychle se vynořila záplava článků, recenzí a natěšených uživatelů. Prvotní nadšení však brzy částečně ochladlo, když se ukázalo, že tento nástroj - jakkoliv pro řadu uživatelů nesporně velmi užitečný, má i mnoho zásadních vad na kráse. Vad, které navíc často přetrvávají až do dnešních dnů. Jak to tedy s Google Scholar vypadá? Je to spící tygr s potenciálem monopolizovat v budoucnu oblast vyhledávání vědeckých informací, obdobně jako klasický vyhledávač Google dominuje dnes obecnému vyhledávání na webu? Nebo

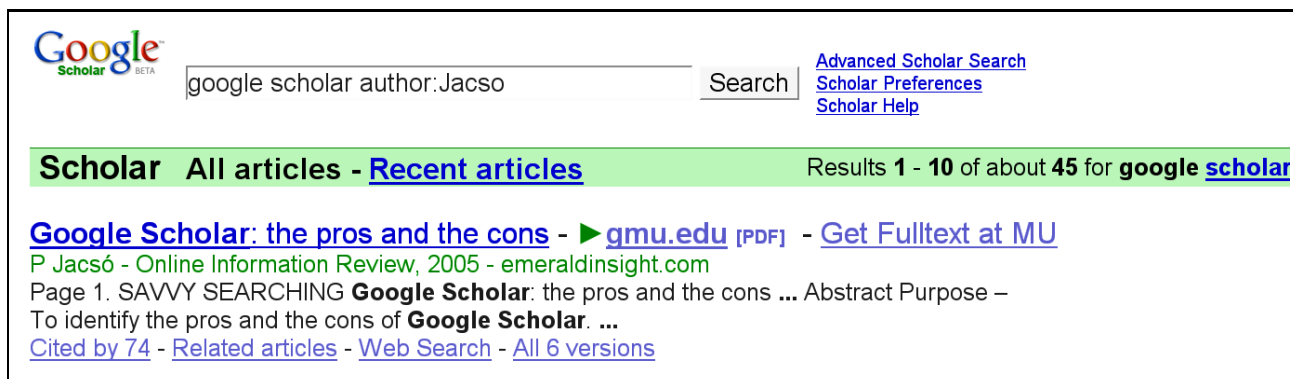
je to ze strany Google spíše již „odpískaný“ nedotažený pokus, slepá ulička vývoje? Těžko hádat, ponechme to budoucímu vývoji. Dnes si alespoň představme Google Scholar jako takový - jeho aktuální stav se všemi obecně uváděnými plusy a mínusy.

1 Co to je

Google Scholar, <http://scholar.google.com>, je služba pro vyhledávání digitálních i fyzických kopií *odborných a vědeckých prací*: recenzovaných časopiseckých a sborníkových článků, preprintů, technických zpráv, odborných knih, vědeckých kvalifikačních prací atd. Cílem je usnadnit uživateli vyhledání potřebné vědecké literatury a nezahrnout ho přitom neodborným balastem, kterým oplývá běžný web. Vyhledávány jsou nejen dokumenty volně dostupné na webu, ale i články v časopiseckých kolekcích akademických nakladatelů, práce v otevřených i uzavřených full-textových, abstraktových a bibliografických databázích, repozitářích učených společností a univerzit, katalogích odborných knihoven. Ne vše, co Google Scholar indexuje a nabízí, musí být nutně (volně) dostupné v plnotextové podobě. Kromě klasického vyhledávání nabízí GS také *citační vyhledávání*, známé ve vědeckém prostředí z profesionálních citačních databází Web of Science či Scopus. Výsledky (plné texty, abstrakty nebo jen bibliografické citace vědeckých prací) jsou sdruženy do skupin odpovídajících jednotlivým dokumentům a jsou seříděny podle relevantnosti a důležitosti vycházející z počtu citací a dalších kritérií.

2 Začínáme s Google Scholar

Práce s GS začíná vyhledáváním - k dispozici je buď klasické jednoduché vyhledávací okénko pro zadání „klíčových“ slov nebo vyhledávací formulář pro přesnější vyhledávání. Obojí je stejné resp. velmi podobné tomu, co známe z klasického vyhledávače Google. Ovšem výsledek vyhledávání již vypadá jinak. Stránka výsledků nezobrazuje přímo jednotlivé dokumenty, nýbrž seznam *prací*. Každá práce je tvořena jedním či několika různými *projevy* téhož díla. Například v případě článků může daný článek existovat



Obrázek 1: Příklad jedné práce v seznamu výsledků Google Scholar

(projevovat se) v podobě preprintu, článku publikovaného v recenzovaném časopise, dokumentu umístěného na osobním webu autora, souboru v univerzitním repozitáři, záznamu v abstraktové databázi nebo pouhé citace. Google Scholar se snaží (automatizovaně) seskupit všechny projevy daného článku dohromady, do jedné vědecké práce. K této práci se pak také vztahuje například počet citací – ten je dán jako součet citací všech projevů dané práce.

Každá práce má klikací název vedoucí na plný text či abstrakt, je-li dostupný on-line. Klikací název nemají samostatně stojící citace (Google Scholar je extrahuje z referencí, bibliografií, poznámek pod čarou apod.) a položky typu kniha.

Pro každou práci nabízí Google Scholar řadu funkcí:

- *Cited by* (Počet citací tohoto článku): uvádí práce, které citovaly danou práci a jsou obsaženy v databázi GS;
- *Related articles* (Související články): hledá další práce, které jsou tematicky podobné článkům dané práce;
- *Web Search* (Hledání na webu): vyhledává informace o dané práci na webu, přes klasický webový vyhledávač Google;
- *All x versions* (Všechny verze – počet: x): uvádí, z kolika projevů se skládá daná práce a umožňuje všechny projevy práce zobrazit.

Kromě těchto základních funkcí mohou být nabídnuty i další funkce, například propojení prací na domovskou instituci uživatele, pokud toto propojení příslušná instituce v GS nastavila. Sem

patří odkaz na projevy práce v katalogu domovské knihovny nebo odkaz na plný text práce v komerční plnotextové databázi, do níž má domovská instituce uživatele přístup (odkaz *Get Fulltext at MU* – viz dále v textu).

3 Přednosti a nedostatky Google Scholar

Google Scholar má své zastánce i odpůrce. Argumenty obou stran se pokusme shrnout v rozboru hlavních předností a nedostatků tohoto nástroje.

3.1 Přednosti

Mezi nejčastěji uváděné přednosti patří obrovský rozsah databáze a její univerzálnost, možnost citačního vyhledávání, které v některých případech poskytuje více výsledků než jiné tradiční nástroje, možnost zapojit do databáze i lokální vědecký obsah instituce, a v neposlední řadě také rychlé vyhledávání, snadnost použití a bezplatné využívání odkudkoliv.

Pokrytí, rozsah

Od svého vzniku prošel největším rozmachem samotný obsah databáze Google Scholar. Ta nyní představuje vůbec nejrozsáhlejší zdroj informací z vědeckého prostředí. Prakticky již každý významný světový producent vědecké literatury je zapojen nebo usiluje o zapojení svého obsahu do Google Scholar. Dokonce i takoví hráči jako Elsevier či Americká chemická společnost, kteří stáli zpočátku bokem, poskytují firmě Google obsah svých komerčních časopisů a databází k indexaci. Kromě zlepšeného časopiseckého pokrytí bylo výrazným impulsem pro rozvoj služby také

zařazení dat z projektu Google Books. Umožnilo zpřístupnit pro vyhledávání plnotextové indexy miliónů digitalizovaných knih. Velkým kladem je také pokrytí značné části digitálních repozitářů a pre/post-printových archivů, zejména v oblastech přírodních věd, medicíny a počítačových věd (byť ne vždy již musí být zaindexován celý repozitář), což zlepšuje mimo jiné i přístup uživatelů k volně dostupné literatuře (open access). Vedle plnotextových zdrojů indexuje Google Scholar také milióny záznamů v bibliografických a abstraktních databázích od předních světových vydavatelů.

Univerzálnost

Google Scholar není specializován na nějakou konkrétní oblast či vědeckou disciplínu, i když jeho pokrytí jednotlivých oborů může být značně rozdílné (mezi nejlépe pokrytými se uvádí přírodní vědy, medicína a počítačové vědy). Široký záběr je velkou výhodou při vyhledávání témat, která se obvykle nevejdou do úzce specializovaných odborných databází, a při multidisciplinárním vyhledávání. Na rozdíl od mnoha jiných odborných zdrojů neomezuje také Google Scholar své geografické a jazykové pokrytí na anglosaský svět, ale indexuje vědeckou literaturu celosvětově a v různých jazycích.

Citační vyhledávání

Jak bylo již zmíněno v úvodní charakteristice, GS nabízí také funkci citačního vyhledávání („kdo cituje můj článek“). Umožňuje tím pohyb informačním prostorem v čase nejen vzad (přes seznam v článku použité literatury), ale i vpřed – od článku k novějším dokumentům, které z daného článku vychází a citují ho. Přestože stávající verze citačního vyhledávání v GS trpí některými neduhy (vyplývajícími hlavně z toho, že veškerá extrakce a analýza citací je prováděna výhradně jen strojovým způsobem, bez dodatečných ručních kontrol a korekcí chyb), pro řadu uživatelů představuje vítaný doplněk klasických komerčních citačních indexů, neboť díky obrovskému záběru databáze GS poskytuje často větší počet citací (byť ne vždy stoprocentně spolehlivých) než specializované citační databáze Web of Science a Scopus.

Napojení na domovskou instituci

V Google Scholar jsou dostupné nejen zdroje od velkých vydavatelů vědecké literatury. I běžné akademické knihovny mohou nabídnout obsah svých databází a katalogů k zařazení do databáze Google Scholar, takže uživatelé mohou být při vyhledávání nabídnuty nejen zdroje na webu či u vydavatelů, ale i odkazy na relevantní literaturu dostupnou v jeho domovské knihovně (služba Library Search). S využitím technologie OpenURL mohou být také výsledky vyhledávání propojeny na plné texty článků v komerčních databázích zakoupených uživatelovou mateřskou institucí (služba Library Link). Tuto funkcionalitu jsme zprovoznili nedávno i pro Masarykovu univerzitu: uživatelé provádějící vyhledávání v GS z počítačů MU uvidí nyní vedle mnoha vyhledávaných položek odkaz „Get Fulltext at MU“, který je zavede přímo na plný text dokumentu v licencovaných fulltextových zdrojích dostupných pro MU.

Google navíc nabízí akademickým institucím digitalizaci jejich vlastních vědeckých časopisů (pokud například celý časopis nebo jeho starší ročníky neexistují v digitální podobě) a jejich začlenění do databáze Google Scholar.

Snadnost použití, rychlost, cena

Podobně jako obecný vyhledávač Google nabízí i Google Scholar velmi snadný způsob zadávání vyhledávacích dotazů, který preferuje velká část uživatelů i v akademickém prostředí. Řada uživatelů nemá chuť učit se složité vyhledávací postupy a studovat sofistikovaná vyhledávací rozhraní. Ve spojení s okamžitou odezvou (Google Scholar využívá propracovaných technologií a zázemí mateřské firmy, které dokáží prohledávat bleskurychle jakkoliv rozsáhlou databázi) může uživatel získat často dostatečně kvalitní odpověď na to, co právě potřebuje. A navíc, vše je zadarmo – na rozdíl od stále rostoucích cen většiny komerčních informačních zdrojů a systémů.

3.2 Nedostatky

Hlavní výtky ke Google Scholar (a je nutno říci, že část informačních specialistů nemá Google Scholar právě v oblibě) spadají do tří oblastí – chyby

softwaru, nedostatečně propracované pokročilé vyhledávání, a zejména nedostatek transparentnosti co se týče obsahu databáze a některých funkcí systému.

Chyby a nedostatky v softwaru

Jeden z nejznámějších kritiků GS, Peter Jacsó, charakterizuje v [1] problém následovně: „zatímco obecný vyhledávač Google odvádí obdivuhodnou práci ve světě nestrukturovaných webových stránek, software Google Scholar pokračuje v mizerných výsledcích při zpracování vysoce strukturovaných a metadaty opatřených vědeckých dokumentů“. Už od uvedení Google Scholar (který vznikl vlastně jako vedlejší produkt resp. experiment v rámci sabbatical období některých vývojářů Google) byly systému vytýkány „školácké“ chyby při analýze strukturovaných vědeckých dokumentů. Například parsovací algoritmy GS nedokáží vždy dostatečně spolehlivě identifikovat u jednotlivých dokumentů klíčové metadatové prvky – jako jména autorů (důsledkem jsou kuriózní autoři typu F. Password nebo V. Chapter), datum (někdy to vypadá, jakoby bylo za datum považováno téměř jakékoliv čtyřmístné číslo v dokumentu, což vede k paradoxům, kdy např. článek z počátku 20. století cituje články ze století 21.) nebo tematické zařazení příslušného dokumentu. To vede k nepřesnostem jednak při pokročilém cíleném vyhledávání (pomocí jmen autorů, data, či s využitím tematických filtrů), zejména ale při citačním vyhledávání, neboť automatické párování citujících a citovaných dokumentů má k dokonalosti dost daleko. (Nutno ovšem přiznat, že vzhledem k chaosu a nezměrné variabilitě panující v oblasti citování, to nemají programátoři GS zrovna snadné). Nespolehlivé a někdy i značně nadsazené bývají kvůli tomu často i odhady počtu článků (počty hitů resp. počty citací).

Zásadní výtky kritiků nesměřují však ani tak k chybám samotným (i v oblasti relativně dobře strukturovaných vědeckých dokumentů je – s ohledem na již zmíněnou variabilitu a obrovský rozsah databáze – automatická extrakce metadat a citací úkol nesmírně obtížný), jako spíše k tomu, že většina chyb přetrvává v systému dlouhodobě. O dosavadní „intenzitě“ vývoje nástroje Google Scholar leccos vypovídá i fakt, že

čtyři roky po svém uvedení je stále ještě označován jako beta-verze.

Nedostatečné pokročilé vyhledávání

Kromě populárního jednoduchého vyhledávání známého z klasického vyhledávače Google nabízí Google Scholar i pokročilé vyhledávání. Uživatel má možnost například omezit vyhledávání na zadané autory (GS ovšem není schopen rozlišit mezi autory stejného jména), časový interval, zdroj (např. časopis) či tematickou kategorii (v české verzi GS není dostupné). Účinnost tohoto vyhledávání je však výrazně poznamenána problémy s ne vždy spolehlivou detekcí metadat (viz předchozí bod) případně s absencí důležitých metadatových prvků u části dokumentů. Navíc jsou možnosti pokročilého vyhledávání GS oproti specializovaným profesionálním databázím poměrně chudobné (což je z větší části daň za univerzální záběr GS). Chybí také možnost vlastního třídění výsledků.

Nedostatek transparentnosti

Zásadní nedostatek však spatřuje velká část kritiků v nedostatku informací o tom, co všechno vlastně Google Scholar indexuje (jaké je jeho pokrytí) a jaký je jeho rating algoritmus. Pro řadu uživatelů by bylo užitečné vědět, jak velká je databáze GS, které časopisy od kterých vydavatelů a za jaké období obsahuje, jaké je přesně pokrytí určité jazykové oblasti, které repozitáře a v jaké míře úplnosti jsou indexovány, jaká je frekvence aktualizace údajů apod. Bohužel, Google Scholar k tomu neposkytuje téměř žádné informace. Většina z toho, co bylo publikováno, vychází jen ze značně nespolehlivých, neúplných či zastaralých experimentů od uživatelů a recenzentů. Chybí i kritérium „vědeckosti“, podle něhož jsou dokumenty do databáze zařazovány (v GS lze najít také editoriály, učebnice, studentské práce a další typy dokumentů, které nemají striktně vědeckou povahu).

Velmi neurčité je rovněž vyjádření vývojářů k tomu, jakým způsobem jsou řazeny výsledky vyhledávání. Je zřejmé, že řazení nemůže být založeno na page-rank algoritmu, jaký používá klasický vyhledávač Google. Zveřejněna byla jen vágní vyjádření v tom smyslu, že při stanovování relevance je do úvahy brána řada faktorů –

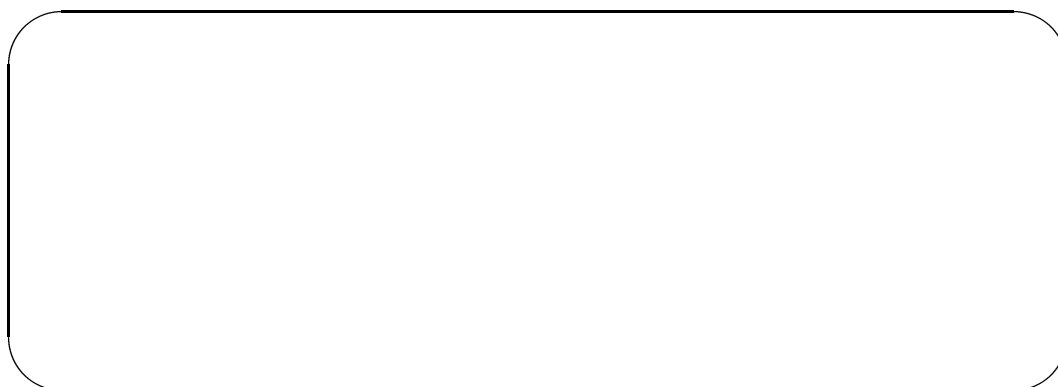
včetně toho, kdo daný článek napsal, kde byl publikován, a především je (nějak) zohledňována citovanost článku. Tato velká míra neurčitosti mj. značně omezuje využití GS pro provádění hodnověrných citačních analýz a porovnání.

4 Shrnutí

Google Scholar je zajímavý a řadou uživatelů často využívaný nástroj pro vyhledávání vědeckých informací. Vzhledem k některým nedostatkům ve svém software však zatím nenaplnuje všechna očekávání vyvolaná při jeho uvedení před čtyřmi roky. Výhodou je obrovská databáze vědeckých informací a její využití zejména při multioborovém vyhledávání. Je neocenitelným pomocníkem pro nalezení rychlé odpovědi a při hledání „jehly v kupce sena“. Pro hlubší a přesnější bádání dává však (zatím?) řada vědců přednost osvědčeným specializovaným oborovým databázím; to platí i pro citační analýzy.

Obsah

Dotazník ke Zpravodaji ÚVT MU – vyhodnocení, Miroslav Bartošek, ÚVT MU	1
Služby ÚVT pro vědu a výzkum, David Antoš, ÚVT MU	2
Velkoformátové tisky na MU, Kamil Malinka, Jiří Ledvinka, ÚVT MU	5
Osm let Celouniverzitní počítačové studovny, Marcela Valentová, ÚVT MU	7
Infrastruktura univerzitních počítačových studoven, Jakub Dobrovolný, Vít Bukač, FI a ÚVT MU	9
Souborný katalog ČR, Hana Vochozková, ÚVT MU	11
Nástroje a služby Google. 2. Google Scholar, Miroslav Bartošek, ÚVT MU	12



Literatura

- [1] Péter Jacsó. *Google Scholar revisited*. Online Information Review, Vol. 32, Iss. 1, 2008. Dostupné online též na <http://www.jacso.info/PDFs/jacso-GS-revisited-OIR-2008-32-1.pdf>
- [2] R. Vine. *Google Scholar*. *Electronic Resources Review*. J Med Lib Assoc, January 2006. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=1324783>
- [3] Tamar Sadeh. *Google Scholar versus Meta-search Systems*. High Energy Physics Libraries Webzine, issue 12, February 2006. <http://library.cern.ch/HEPLW/12/papers/1/>
- [4] *An interview with Anurag Acharya, Google Scholar lead engineer*. Google Librarian Central - Article 12/2006. http://www.google.com/librariancenter/articles/0612_01.html □