

Vyhledávací server Novell QuickFinder Server

Když jsme před deseti a více lety potřebovali získat nějakou informaci, sáhli jsme například po telefonu, denním tisku, odborném časopisu, encyklopedii apod. Nyní to máme mnohem jednodušší, protože existuje Internet. Je zde však jeden problém - v tom obrovském množství informací není jednoduché tu požadovanou rychle najít. Zde nám pomáhají především webové vyhledávací služby (např. Google). Jedním z výkonných prostředků umožňujících realizovat služby tohoto typu, je i produkt QuickFinder Server od společnosti Novell.

QuickFinder Server je špičkový vyhledávací prostředek. Může směle soutěžit s nejlepšími vyhledávacími službami od IBM, Microsoftu, Googlu, AltaVisty apod. Mnoho zákazníků společnosti Novell, kteří používají její produkty a služby, o něm však vůbec neví. Přitom se s ním setkávají, kdykoli něco hledají na webových stránkách novell.com. Navíc řada z nich, aniž si to uvědomuje, ho možná už vlastní a mohla by jej tedy bezplatně využívat.

Novell dal tento prostředek k dispozici již ve svém síťovém operačním systému NetWare 5.1, a to pod jménem *NetWare Web Search Server*. Následně jej vyvíjel a zařadil i do NetWare 6 a 6.5. Jeho nové jméno *QuickFinder Server* pak vyjadřuje především skutečnost, že je dodáván v rámci produktu Novell Open Enterprise Server (NOES), přičemž je tam k dispozici nejen pro platformu NetWare, ale i SUSE Linux.

Informace uvedené v tomto článku jsou zajímavé především pro ty, kdo ještě nemají ve své organizaci vyhledávací službu instalovanou, kteří ji mají instalovanou ale často nemohou požadované informace najít, jejichž vyhledávací služba indexuje webové servery ale neumí indexovat běžné soubory, a také pro ty, kdo vlastní některý ze zmíněných systémů NetWare či NOES a přitom zbytečně někomu platí za poskytování vyhledávacích služeb. Jsou zde uvedeny jak informace o samotném QuickFinder Serveru, tak i o způsobu jeho realizace na novell.com.

Koncepce

QuickFinder Server umožňuje uživatelům vyhledávat požadované informace na veřejných i privátních webových serverech své organizace, na jejích souborových serverech i na cizích webech dostupných přes Internet.

Základním pojmem souvisejícím s QuickFinder Serverem je **virtuální vyhledávací server** (virtual search server). QuickFinder Server je jich schopen realizovat na jednom hostitelském serveru libovolný počet. Každý z nich je plně nezávislým vyhledávacím místem sloužícím dané organizaci, oddělení či webovému serveru. Každý má své vlastní indexy, konfigurační nastavení, naplánované události, vzory pro zobrazování výsledků, rozhraní pro správu atd.

Z obsahu zadaných webových či souborových serverů generuje QuickFinder Server **indexy**, tzn. soubory, jež obsahují vyhledávaná klíčová slova a URL-adresy odpovídajících dokumentů. Každý virtuální vyhledávací server jich může mít několik. Jejich prostřednictvím se pak vytváří v reakci na dotazy uživatelů seznamy vyhledaných dokumentů. Lze přitom vytvářet dva typy indexů: *Crawled* a *File system*. První indexuje obsah webových serverů, druhý pak serverů souborových.

Významnou součástí QuickFinder Serveru jsou i **vzory pro zobrazování výsledků** (templates). Jejich prostřednictvím je řízeno rozhraní, přes něž se předávají výsledky vyhledávání uživatelům. Díky nim lze výsledkové stránky snadno modifikovat a dosáhnout tak jejich požadovaného vzhledu. Jednotlivé virtuální servery mohou tyto vzory vzájemně sdílet nebo každý z nich může mít definovanou vlastní sadu vzorů. Jsou k dispozici ve formátech html a xml.

Poskytované funkce

QuickFinder Server disponuje všemi vlastnostmi a funkcemi, které se u moderního vyhledávacího prostředku dají očekávat. Alespoň pro představu se nyní v krátkosti podívejme na charakteristiku některých z nich:

- *Typy dokumentů*. Jsou podporovány více než dvě desítky souborových formátů, např. html, xml, pdf, Word, WordPerfect, Excel, PowerPoint, OpenOffice apod.

- *Národní jazyky*. QuickFinder Server podporuje dokumenty napsané ve více než dvou desítkách různých jazyků (např. v angličtině, francouzštině, němčině, ruštině, čínštině, japonštině, polštině, řečtině, arabštině, hebrejštině) a rozeznává více než 50 znakových sad.
- *Stanovení významnosti* (Relevance Adjustments). Pořadí položek v zobrazeném seznamu výsledků vyhledávání je obvykle dáno výskytem hledaných výrazů v příslušných místech dokumentů (v příkazech Title, Description, hlavičkách apod.). QuickFinder Server však umožňuje stanovit významnost dokumentů i uměle. Správce vyhledávací služby tak může zvyšovat význam jednotlivých dokumentů, dokumentů ze zadaných webů či jejich sekcí, z určitých indexů apod.
- *Nejlepší tipy* (Best Bets). Toto je funkce, jejímž prostřednictvím lze upozornit uživatele na významné dokumenty. Umožňuje poskytovat kromě normálního seznamu výsledků i seznam nejlepších tipů. Ten je pak zobrazován jako první, přičemž je vytvářen z indexů, jež navrhuje správce speciálně pro tento účel. Tímto způsobem se např. upozorňuje na informačně nejhodnotnější dokumenty, nejnovější nabídky, zvláště doporučované produkty apod.
- *Zvýrazňování* (Highlight). Díky této funkci jsou v zobrazených dokumentech zvýrazňována slova, jež uživatelé zadali k vyhledání.
- *Duplicitní URL*. Technologie QuickFinder Serveru označovaná „fingerprint“ detekuje a odstraňuje při indexování většinu duplicitních URL-adres.
- *Odkazy na jiná webová místa* (Off-Site Links). Jsou-li při indexování webového serveru nalezeny v jeho dokumentech odkazy mířící mimo zadaný prostor URL-adres, jsou normálně přeskočeny. Zde uváděná funkce však umožňuje požadovat, aby byla zpracována i zadaná sada „vnějších“ URL-adres. Toho se využívá např. pro odkazy mířící na stránky cizích webů, které se týkají dané organizace (články o organizaci, jejich produktech, weby partnerů apod.).
- *Nežádoucí soubory* (Unwanted files). Tato funkce umožňuje vynechávat dokumenty, jež jsou nežádoucí. Umí vynechat i zbytečné části dokumentů, např. společné hlavičky, zápatí apod.
- *Kontrola správnosti psaní* (Speller) - Pokud výraz zadaný uživatelem pro vyhledání vygeneruje příliš málo výsledků, QuickFinder Server díky této funkci automaticky doporučí alternativní výrazy, jež má definovány ve svém slovníku. Tímto způsobem je vhodné nabízet např. slova, která uživatelé často uvádějí nesprávně.
- *Synonyma*. Pomocí seznamu synonym lze zařídit, aby se do zadaného dotazu automaticky vkládala slova, jež mají podobný význam, jako výrazy zadané uživatelem k vyhledání (např. program - aplikace - produkt).
- *Bezvýznamná slova* (Stop words). Tato funkce odstraňuje z dotazů slova, která mají omezený nebo nulový význam (např. spojky a předložky).
- *Klastry vyhledávacích serverů* (Clustering). Z vyhledávacích serverů lze vytvářet skupiny (klastry). Jejich konfigurační nastavení, nově vygenerované indexy a vyhledávací vzory jsou přitom automaticky vzájemně synchronizovány. V rámci správy celé skupiny serverů pak tedy stačí konfigurovat jen jeden server klastru.
- *Záznamy o dotazech* (Query Reports). QuickFinder Server udržuje detailní informace o provedených vyhledávacích operacích. Díky tomu mohou správcové získávat detailní týdenní a měsíční záznamy o činnosti této služby. Logové soubory lze přenášet i na pokročilé prostředky pro zpracování reportů a provádět tam jejich analýzu.
- *E-maily*. QuickFinder Server se snadno integruje s elektronickou poštou, což např. umožňuje zasílat správcům chybové zprávy, varování, záznamy o dotazech apod.

Instalace, použití a správa

QuickFinder Server se **instaluje** velmi snadno. Při standardním generování systému NOES se o něj dá požádat jako o volitelnou součást nebo jej lze instalovat na vyhrazený server použitím příslušného instalačního vzoru. Při instalaci vznikne na hostitelském serveru vzorový virtuální vyhledávací server a ten má vytvořeny dva indexy: *QuickFinder Documentation* pro soubory nápovědy tohoto produktu a *DocRoot* pro obsah aktuálního webového serveru (tzn. dostupného přes port 80). Po instalaci je možné využívat vzorový server nebo vytvořit další virtuální vyhledávací servery. Poté se definují a generují příslušné indexy, zadávají požadované parametry a případně se modifikuje podle vlastních potřeb sada vzorů pro zobrazování výsledků.

Jednoduché je i **použití** zmiňované vyhledávací služby. Uživatel si otevře ve svém webovém prohlížeči úvodní stránku vyhledávacího serveru (např. na <http://server/qfsearch/search>) a zde zadá svůj požadavek (obr.1). Vyhledávací server pak prohlídne příslušný index a hledá v něm zadaný výraz. Pokud najde odpovídající dokumenty, zobrazí uživateli prostřednictvím vzoru jejich seznam (obr.2).

Ke **správě** zmiňovaného vyhledávacího prostředku slouží webový nástroj *QuickFinder Server Manager* (obr.3). Je dosažitelný přes iManager a jeho odkaz QuickFinder nebo na adrese <http://server/qfsearch/admin>.

Prostřednictvím tohoto nástroje se spravují jednotlivé virtuální vyhledávací servery, jejich indexy, parametry, logové soubory apod.

Realizace na novell.com

Jako příklad konkrétní implementace QuickFinder Serveru, s nímž se můžeme běžně setkat, lze uvést webové místo *novell.com*. Vyhledávací služba je zde poměrně značně zatížená, řeší kolem 4 milionů dotazů měsíčně. Proto Novell vyhradil k tomuto účelu 4 servery, jež zajišťující vyhledávání pro novell.com a fóra Novellu a 2 servery pro interní web Novellu. Dále jsou použity 2 servery, které denně regenerují každý index. Všechny 6 vyhledávacích serverů je zpřístupněno okolnímu světu přes iChain-server, který mimo jiné zajišťuje vhodné rozložení zátěže.

Zmíněné vyhledávací servery Novellu mají momentálně následující parametry: 2 procesory 1.2 GHz, paměť 2 GB a disky 34 GB. Každý indexovací server je vybaven dvěma procesory 800MHz, pamětí 1 GB a disky 17 GB. Díky tomu, že servery zajišťují jen zmíněné služby, zvládají vysokou zátěž, i když jejich výkonové parametry už nejsou špičkové. Na uvedených serverech se využívají systémy Open Enterprise Server SP 1 a SUSE Linux Professional 9.2.

Všechny funkce, jež poskytuje vyhledávací služba na novell.com, mohou mít pochopitelně k dispozici ve svých sítích (nebo už mají) i zákazníci Novellu, kteří již QuickFinder Server vlastní.

Ing. Oldřich Přichystal

Zpracováno volně podle článku „Lost and Found“ uveřejněného v Novell Connection Magazine 11+12/2005.

Legenda k obrázkům:

QFIND_O1 Obr.1 Okno QuickFinder Serveru pro vyhledávání

QFIND_O2 Obr.2 Okno s výsledky vyhledávání

QFIND_O3 Obr.3 Okno QuickFinder Server Manageru