

Mit DVD Top-Programme für Freizeit & Beruf € 9,90 November-Jahrgang 1/07

MAGAZIN FÜR FREIE SOFTWARE

OpenSource

4 197-1306-30393701

Projekte zum Mitmachen:

- Linux im Eigenbau**
Schritt für Schritt zur ganz persönlichen Distribution
+ alle nötigen Tools auf Heft-DVD
- Google-Alternative**
So helfen Sie beim Aufbau einer vernetzten Suchmaschine
+ Software auf Heft-DVD
- Gratis fliegen lernen!**
Kostenloser Simulator mit zahlreichen Flugzeugtypen
+ XP-Software auf Heft-DVD
- Photoshop-Turbo**
Extrafeatures für Fototools
+ Add-ons auf Heft-DVD
- USB-Stick-Tools**
Clevere Zusatzfunktionen:
PC-Schlüssel, Datentresor ...
+ Toolbox auf Heft-DVD

1.800 **4,7 GB gratis auf DVD**

aktuelle Programme für Windows und Linux

OpenSource

- Ubuntu Linux 6.10 RC**
Komplette Linux-Distribution für Umsteiger und Fans
- XP-Turbo-Tuner**
100 Windows-Optionen in nur einem Fenster einstellen
- Dark Fox 2.0**
Böser Firefox-Bruder mit genialen Download-Plug-ins

OpenOffice.org Premium

90 Schriften
Dokumentvorlagen
Einfache Bedienung
Gratis-Updates

inklusive **2.800 Cliparts**

+ 30 deutsche Vollversionen im Wert von über 800 Euro*

PCPro@X-Sonderheft

YaCy im OpenSource Magazin 1/07:

- Das Projekt wird als „Google-Alternative“ bezeichnet und erscheint als solches auf dem Titelblatt
- Die „Google-Alternative“ wird im Editorial erwähnt
- YaCy ist Haupt-Aufmacher im Inhaltsverzeichnis
- 4 Seiten Artikel mit Installationsbeschreibung und Dokumentation

OpenSource Editorial

Mitmach-Projekte

Was nichts kostet, taugt auch nichts: Dieser Spruch lässt sich auf Open Source mit Sicherheit nicht anwenden. Schließlich hat die Community eine Vielzahl Programme entwickelt, die es durchaus mit kommerziellen Produkten aufnehmen können. Wir haben für Sie einige dieser Highlights auf unserer Heft-DVD zusammengetragen: vom kompletten Betriebssystem bis zum piffigen XP-Tuning-Tool. Doch Open-Source-Software ist eben nicht nur kostenlos wie Freeware, sie ist offen und frei. Das heißt, dass Sie die Programme nicht nur verwenden, sondern auch anpassen oder erweitern dürfen. Damit Sie selbst loslegen können, präsentieren wir Ihnen im Heft zahlreiche Mitmach-Projekte, von einer Google-Alternative, über Firefox-Tuning bis zum Eigenbau-Linux. Kennen auch Sie ein spannendes Open-Source-Projekt, das wir im kommenden Heft einmal vorstellen sollten? Dann schreiben Sie mir mit dem Betreff „Mitmach-Projekt“ an adolny@databecker.de.



Dr. M. Dolny
Dr. Michael Dolny,
Chefredakteur

MAGAZIN FÜR FREIE SOFTWARE

OpenSource

Inhalt 1/07

Magazin

- 6 News aus der Open-Source-Szene
- 22 Treibertrouble Closed-Source-Treiber für Linux
- 25 Der Überall-Pinguin Linux auf Embedded Devices
- 28 Das Open Source Jahrbuch 2006 Großes Nachschlagewerk
- 30 Desktop marsch Linux als Unternehmens-Client

Top-Programme

- 36 WinOptimizer 2007 XP optimal einstellen
- 44 Daten perfekt abgleichen SyncroNaut - Datei kopieren deluxe
- 46 Burning Studio 2007 Die Brennseite für alle Fälle
- 48 Kostenloser Komplettenschutz Maximale Sicherheit im Gratispaket
- 52 Die größte Tool-Sammlung aller Zeiten Gratis-Tool-Archiv
- 57 Gratis fliegen lernen Der kostenlose Flugsimulator FlightGear
- 61 OpenOffice 2 Premium Kostenloses Komplettpaket fürs Büro

Windows-Projekte

- 64 Exklusiv: XP-Turbo-Tuner Windows perfekt optimiert
- 68 Sicherheitstuning Maximaler Schutz mit dem XP-Turbo-Tuner
- 70 Dienste im Griff Windows mit dem Turbo-Tuner den letzten Schliff geben
- 72 Windows-Desktop Microsoft-frei Microsoft nur im Untergrund
- 75 Die Toolbox für Ihren USB-Stick fast quadratisch, praktisch, gut

Internet-Projekte

- 78 „Vista“-Fox Der beste Browser für das neue Windows
- 81 Dark Fox 2.0 Firefox, jetzt noch böser
- 84 Pimp my Browser Die besten Plug-Ins für Firefox
- 94 Gemeinsam gegen Google Alternative Suchmaschine

Linux-Projekte

- 98 Mini-Linux im Eigenbau Schritt für Schritte zur eigenen Distribution
- 102 Pinguin serviert Datei- und Druckserver mit Ubuntu Linux bauen
- 106 Experimentierfreude Ubuntu 6.10 und openSUSE 10.2
- 110 Durchstarten mit SUSE Linux Remastered
- 114 Desktops neue Kleider Windowmanager jenseits von KDE und GNOME
- 118 Erstkontakt Einstieg in die Linux-Shell

Foto, Musik & Video

- 122 Open Source für Fotografen kostenlose Profi-Tools
- 128 Photoshop-Turbo Leistungsschub für Photoshop & Co.
- 136 Besser und freier Musik genießen Codex & Co.
- 140 Fotos optimieren Profi-Bildbearbeitung mit GIMP 2.3

Service

- 3 Editorial
- 146 Impressum
- 5 DVD-Inhalt
- 146 Vorschau

Die Toolbox für Ihren USB-Stick

Zum Transport von Daten und Dokumenten sind USB-Sticks eine praktische Sache. Die Mini-Datenträger können aber mehr: Ab Seite 75 präsentieren wir Ihnen Tools, mit denen Sie Ihren USB-Stick in ein geniales Multifunktionsgerät verwandeln.

Gratis fliegen lernen!

FlightGear ist einer der attraktivsten Flugsimulatoren und vor allen Dingen kostenlos erhältlich. Er steht unter der GPL und läuft unter den Betriebssystemen Windows, Linux, Mac OS X, Solaris und FreeBSD. Wir zeigen ab Seite 57, wie Sie mit FlightGear abheben.

Google-Alternative

Es muss nicht immer Google sein. Als Kombination aus Peer-to-Peer und Distributed Computing möchte die freie Suchmaschine YaCy den Platzhirschen Google & Co. die Stirn bieten. Wir erklären Ihnen ab Seite 94, wie Sie mit Ihrem PC quasi nebensächlich beim Aufbau einer Google-Alternative helfen können.

Photoshop-Turbo

Neue Effekte, Filter und Makros. Für Photoshop und viele andere Bildbearbeitungsprogramme gibt es eine riesige Auswahl Zusatzfunktionen. Wir haben die interessantesten Erweiterungen aufgespürt und zeigen Ihnen ab Seite 128, wie Sie mit dem kostenlosen Tool alles anpassen können.

OpenOffice 2 Premium

Sie suchen ein leistungsfähiges und vor allem kostenloses Office-Paket? Dann greifen Sie doch zum OpenOffice! Die Premium-Version der Suite finden Sie komplett auf unserer Heft-DVD. Außerdem erfahren Sie ab Seite 61, was diese neue Version außer 2.800 Cliparts, 90 Schriftarten und vielen Vorlagen noch alles zu bieten hat.

OpenOffice.org Premium

90 Schriften
Dokumentvorlagen
Sicherer als Word

DVD-Inhalt



Programme auf der Heft-DVD

30 Vollversionen
Eine große Anzahl an kostenlosen Vollversionen für die verschiedensten Bereiche finden Sie auf der Heft-DVD:

- Windows optimieren**
Mit dem WinOptimizer 2007 gelangen Sie in wenigen Minuten zu einem schnellen, stabilen und sicheren Windows XP. Die Systempflege-Suite bietet eine Vielzahl nützlicher Features.
- Sicherungskopien erzeugen**
VersionBackup 4.0 sichert Ihre Daten optimal. Die Backups werden mit Datum abgespeichert und lassen sich - falls nötig - über einen speziellen Backupbrowser finden und wiederherstellen.
- Open Source**
Ob im Office-, Grafik- oder Business-Bereich: Zu den meisten Windows-Programmen finden Sie eine kostenlose Alternative. Neben OpenOffice.org 2 Premium, GIMPShop mit Photoshop-Funktionalität oder dem komfortablen Webeditor oder Nvu finden Sie noch eine Reihe weiterer nützlicher Tools auf der DVD.
- 200 Freeware-Programme**
Das große Gratis-Tool-Archiv mit 200 Freeware-Programmen auf unserer Heft-DVD lässt kaum noch Wünsche offen. Neben den besten Tuning-Tools für Windows XP bekommen Sie alle Werkzeuge für die Bearbeitung von Musik, Fotos und Videos.
- USB-Stick-Tools**
Ein USB-Stick eignet sich nicht nur hervorragend für den Transport von Daten und Dokumenten, der Mini-Datenträger lässt sich auch in ein Multifunktionsgerät verwandeln. Auf unserer Heft-DVD finden Sie nützliche Tools, mit denen Sie Ihren USB-Stick belegen können.

OpenOffice.org Premium

90 Schriften
Dokumentvorlagen
Sicherer als Word

Linux-Anwendungen

Auf der Heft-DVD haben wir für Sie die ersten drei ISOs der **SUSE Linux 10.1 Remastered** vom 9. Oktober 2006 bereitgestellt. Die SUSE-Remastered ist eine stabile Distribution mit einer umfangreichen Softwareausstattung, die sich mit Windows messen kann und sich problemlos neben einem bestehenden XP installieren lässt. **Ubuntu 6.10 RC** basiert auf der Distribution Debian und ist auf Benutzerfreundlichkeit ausgelegt. Im Heft erläutern wir, wie Sie mit Ubuntu und Xming einen Datei- und Druckserver bauen.

Gemeinsam gegen Google

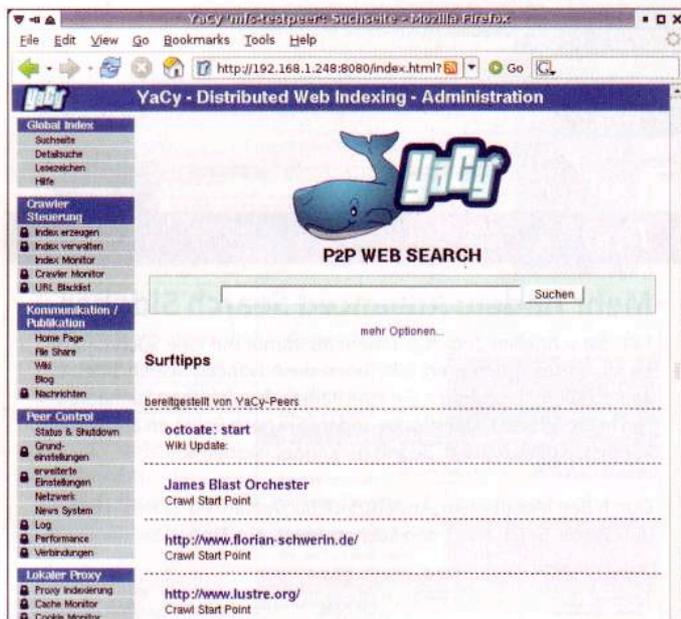
Als Kombination aus P2P-Technologien und Distributed Computing möchte die freie Suchmaschine YaCy Google und Co. die Stirn bieten. Abwegig ist dieser Plan nicht – stehen doch einigen zehntausend Rechnern in Googles Cluster Millionen potenzieller YaCy-Peers gegenüber (Mattias Schlenker/md)

Das Prinzip von YaCy klingt zu einfach, um wahr zu sein: In jedem kleineren Netzwerk dient ein Rechner als Knoten im Suchmaschinennetz. Dieser Knoten kann Seiten ganz klassisch indexieren, indem er als Spider oder Crawler über Teile des WWW geschickt wird,

angesurfte Seiten nicht nur zwischenspeichert, sondern auch indiziert – und auf Wunsch von diesen Seiten aus verlinkte Seiten gleich dazu.

YaCy gegen Google 1:0

Gerade diese Indexierung tatsächlich angesurfter Seiten kön-



Neben der Suchmaske zeigt die YaCy-Startseite gemeinsame Bookmarks und Startpunkte von Webcrawls

von wo aus er mit gesammelten Daten zurückkommt. Indexierte Seiten sind zunächst lokal verfügbar und werden – wenn der YaCy-Port 8080 nach außen freigegeben ist – auch anderen YaCy-Peers zur Verfügung gestellt. YaCys Standard-Suchmethode ist jedoch pfiffiger: YaCy läuft im Hintergrund als Proxy mit, der

nen Suchmaschinen wie Google nicht bieten. Google und Co. ziehen als Kriterium für die Relevanz einer Seite vor allem die Anzahl der auf eine Seite verweisenden Links sowie deren „Gewicht“ heran. Ein Link von einer hoch gerankten Seite ist demnach für das Ranking der verlinkten Seite oft mehr wert als



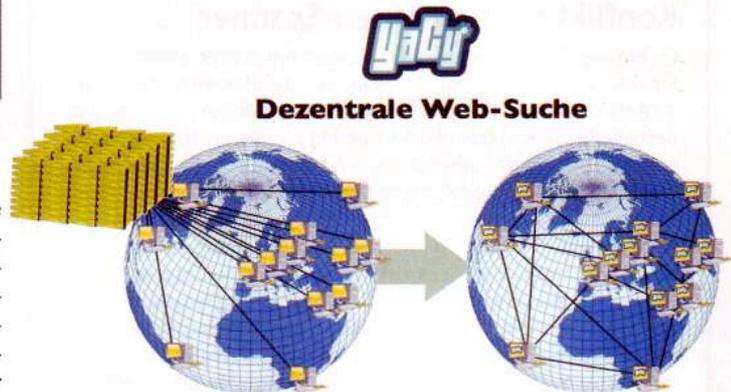
mehrere Links von schlechter gerankten Sites. Doch ein Kriterium kann Google nicht erfassen: wie viele Surfer eine bestimmte Seite tatsächlich aufsuchen.

YaCy kann tatsächliche Besuche erfassen und diese ins Ranking einfließen lassen. Häufig angesurfte Sites sind so aktueller im Cache und im Index enthalten und werden von mehr YaCy-Knoten vorgehalten, was die Wahrscheinlichkeit, diese Site schnell zu finden, deutlich erhöht. Die großen Suchmaschinenbetreiber sind sich des YaCy-Vorsprungs hinsichtlich tatsächlich besuchter Seiten durchaus bewusst. Sie arbeiten deshalb teilweise an Browser-Erweiterungen, die auch Listen tatsächlich besuchter Seiten (ohne personalisierte Seiten) an die Suchmaschinenbetreiber übermitteln. Ein Beispiel für solch eine Erweiterung ist Googles Web-accelerator, der als Tausch für die übermittelten URL-Listen einen vorausschauenden lokalen Cache bereitstellt – nicht oh-

ne in der Vergangenheit durch teils massive Sicherheitslücken aufzufallen.

Dunkle Ecken finden

Schätzungsweise ein Drittel aller Seiten im WWW sind von den Suchmaschinen schlecht bis gar nicht erfasst und damit dem „Deep Web“ zuzuordnen. Mögliche Ursachen sind fehlende Links auf gut gerankten Seiten oder es handelt sich um Seiten, deren URLs eher in Foren oder per Mail ausgetauscht werden. Es muss sich hierbei keinesfalls um illegale Seiten handeln: Viele normale private Seiten, einige Howtos und so manche Blogs sind bei den großen Suchmaschinen einfach in die Bedeutungslosigkeit abgerutscht, weil die Aktualisierungsfrequenz niedrig war. Da YaCy primär besuchte Seiten indiziert, dringt die P2P-Alternative auch in die dunklen Ecken vor. Allerdings gelten jenseits bereits durch den Cache indizierter



Statt eines zentralen Indexes pflegt jeder Knoten einen kleinen Index und tauscht diesen mit anderen Hosts aus

So geht's

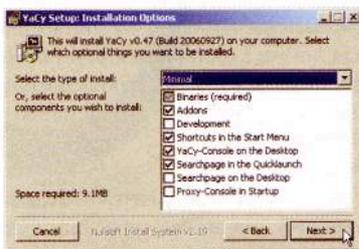
Installation von YaCy

Man könnte denken, dass eine umfangreiche, Java basierte Serversoftware eine aufwändige Installation zur Folge hat. Das Gegenteil ist der Fall: Eine Basisinstallation ist mit wenigen Klicks erledigt.

1. Auf Ihrem System muss eine aktuelle Java-Laufzeitumgebung von Sun vorhanden sein. Wahlweise kann Version 1.4.2 oder 1.5.0 zum Einsatz kommen. Sollte eine ältere Version oder gar Microsofts JVM installiert sein, installieren Sie Java 1.5.0 von der Heft-DVD. Alle vorgeschlagenen Einstellungen können Sie in der Regel übernehmen.



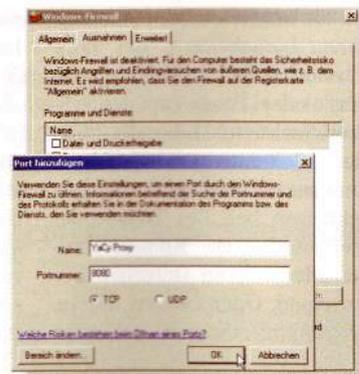
2. Die Windows-Version von YaCy ist als komfortabler Installer in EXE-Form ausgeführt und enthält alle notwendigen Batch-Dateien und Skripte für den Betrieb des Suchmaschinenproxies. Wählen Sie die vorgeschlagene Minimalinstallation. Diese genügt für die Nutzung als Senior- oder Junior-Peer. Zusätzliche Komponenten werden nur benötigt, wenn Sie an YaCy mitprogrammieren wollen.



3. Ein Doppelklick auf den YaCy-Starter aktiviert den Proxy und holt den Webbrowser in den Vordergrund. Hier ist die Grundkonfiguration von YaCy vorzunehmen: Wählen Sie ein sicheres Passwort und lassen Sie sich einen sinnvollen Namen des Peers einfallen. Der Port kann meist auf 8080 belassen werden.



4. Soll Ihr YaCy-Knoten seinen Index mit anderen Nodes teilen – was notwendig ist, um vom Junior zum Senior aufzusteigen –, müssen Sie eingehende Verbindungen auf Port 8080 in Ihrer Firewall zulassen.



5. Falls Sie per DSL- oder Kabelrouter ins Internet gehen, müssen Sie auch hier ein Portforwarding zum YaCy-Host in Ihrem Netzwerk aktivieren. Wenn Ihre Rechner ihre IP-Adressen per DHCP beziehen, finden Sie sie mit dem Befehl

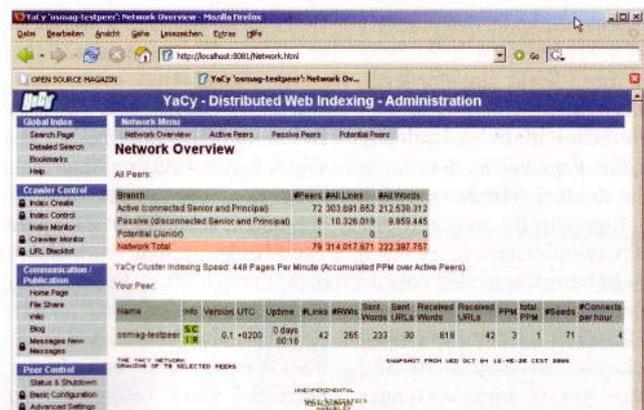
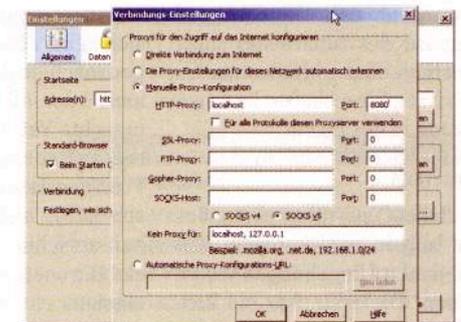
```
ipconfig /all
```

in der Windows-Eingabeaufforderung, die Adresse des Rechners heraus.

6. Stellen Sie in den Verbindungseinstellungen Ihres Browsers als HTTP-Proxy „localhost“ mit Port „8080“ ein. Unter Firefox finden Sie diese Option unter „Extras/Einstellungen/Verbindungseinstellungen“. Auch von anderen Rechnern im lokalen Netz können Sie mit dem Proxy verbinden. In diesem Fall müssen Sie statt „localhost“ die IP-Adresse (beispielsweise 192.168.1.50) der Netzwerkkarte Ihres YaCy-Rechners einstellen.



7. Surfen Sie ein wenig im Netz und werfen Sie anschließend einen Blick auf die Netzwerkwarte von YaCys Peer-Konfiguration (<http://localhost:8080/Network.html>). Jetzt sollte Ihr Knoten im Info-Feld „SCIR“ stehen haben. Der erste Buchstabe bestätigt, dass Sie nun als Senior-Peer zu YaCy beitragen.

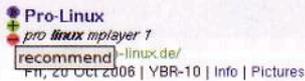


Seiten die Beschränkungen aller Suchmaschinen: Durch POST-Parameter aufgerufene Seiten werden überhaupt nicht und via GET-Parameter angesurft Seiten beschränkt indiziert.

YaCy im Intranet
Prinzipiell ist es Yacy egal, in welchem Netz er indexieren und crawlen darf. Die Eigenschaft können sich Unternehmen zu-

nutze machen, die YaCy im Intranet einsetzen: Richtet man den Proxy so ein, dass er seine Indexes nicht nach außen übermittelt, wird aus der P2P-Suchmaschine eine Intranet-Such-

maschine, die dennoch Ergebnisse anderer Yacy-Knoten mitbenutzen kann. Trotzdem kann die Proxy-Funktionalität aktiv bleiben, was einerseits erlaubt, das Surfverhalten der Mitarbei-



Die Komponente Mensch: YaCy erlaubt die Bewertung von Seiten und deren Aufnahme in eine globale Lesezeichenliste

ter statistisch zu erfassen, andererseits können unerwünschte Seiten auf eine Blacklist gesetzt und damit gesperrt werden. Da angesurft Seiten indexiert werden, ist zudem bei Firmen, die in speziellen Branchen arbeiten, die Wahrscheinlichkeit höher, dass im Cache relevante Suchergebnisse gefunden werden. Insgesamt bietet YaCy für den Einsatz im Firmennetz einen guten Kompromiss aus Suchmaschinen- und Proxy-Funktionalität und dem berechtigten Interesse der Mitarbeiter nach Wahrung eines Mindestmaßes an Privatsphäre bei der Internetnutzung in der Firma.

In der Praxis

Wir haben YaCy über ein halbes Jahr lang in zwei komplett verschiedenen Umgebungen testen können. Die erste war ein kleines Büronetz mit zwei Arbeitsplätzen, an denen Internetrecherche betrieben werden konnte. Testumgebung zwei waren die Netzwerke von zwei Jugendtreffs, die mit jeweils sieben Clients und massivem Netzwerkverkehr zu Stoßzeiten als Belastungsprobe für YaCy gelten durften. Im ersten Netzwerk konnte die Proxy-Indexiertiefe auf 1 gesetzt werden – das heißt, dass nicht nur direkt angeforderte, sondern auch einen Link entfernte Seiten im Index landeten, was die Ergebnisse der YaCy-Suche deutlich verbesserte. Bei den Jugendtreffs mussten wir die Proxy-Indexiertiefe auf 0 setzen – höhere Werte sind sowohl an der entstehenden Netzwerklast als auch an der Systemlast durch die ca. 20 bis 30 Mal höhere Anzahl auszuwertender

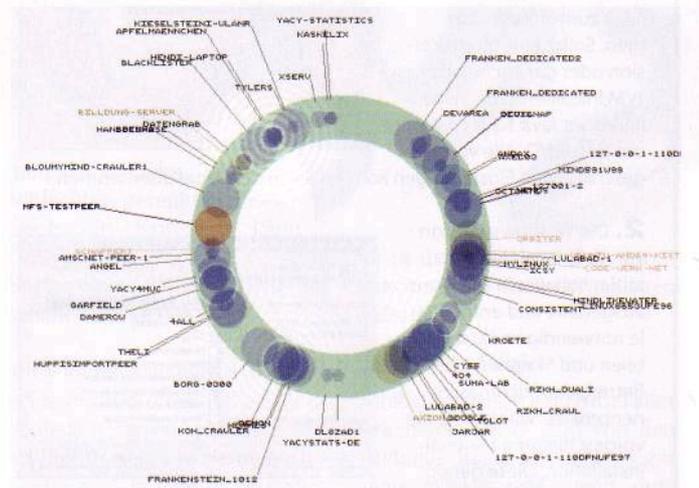
Seiten gescheitert. Alle Testsysteme liefen sieben Tage pro Woche und 24 Stunden pro Tag durch, um nachts möglicherweise aufgelaufene Warteschlangen abzuarbeiten und anderen Peers für so genannte Remote-Crawls zur Verfügung zu stehen. Mit YaCy-Versionen kleiner als 0.47 hatten wir hier jedoch alle paar Tage das Problem, dass der Proxy nicht erreichbar war: YaCy scheiterte an Speicherlimits und musste neu gestartet werden. Wir lösten das Problem schließlich mit der Holzhammermethode, indem wir einen morgendlichen Neustart erzwangen. Neuere YaCy-Versionen gehen etwas behutsamer mit dem Speicher um – den täglichen Neustart behielten wir dennoch bei. Hinsichtlich Einsatz auf älterer Hardware sollten Sie sich keine allzu großen Hoffnungen machen: Auf allen Testsystemen der 2-GHz-Klasse verursachte YaCy im normalen Einsatz zwischen 20 und 60 Prozent Prozessorlast und zwackte zwischen 100 und 300 MB des Arbeitsspeichers ab. Als Proxy für ein Büronetz sollte deshalb wenigstens ein Rechner der Gigahertz-Klasse mit 256 MB RAM zum Einsatz kommen. Bei Nutzung auf einem Einzelplatzsystem bleibt der P2P-Proxy erst unbemerkt, wenn etwa 512 MB RAM und ein Prozessor im 1,5-GHz-Bereich eingebaut sind. Eine Nutzung auf NAS-Geräten wie der NSLU2 oder erweiterbaren Routern wie der FRITZ!-BOX dürfte in naher Zukunft also nicht in Frage kommen.

Begehrlichkeiten

Der autonome – oder besser gesagt autarke – Ansatz ohne zentrale Kontrolle weckt Begehrlichkeiten, schließlich kann jede YaCy-Node einen prinzipiell gleichberechtigt behandelten Crawl starten, an dem andere Knoten mitwirken. Crawl-Startpunkte erscheinen auch auf der Suchseite jeder YaCy-Node als

„Surftipp“. Doch die Gefahr, dass Spam und politischer Extremismus im YaCy-Netz überhand nehmen, ist gering. Zum einen, weil jeder Administrator bestimmte Seiten von Crawl und Indexierung ausschließen kann – und sich die Blacklists einfach austauschen lassen –, zum anderen, weil Anwender „Surftipps“ und Suchergebnisse positiv oder negativ bewerten können. So

gebnisliste oft gut bis befriedigend. Schlechter sieht es beispielsweise bei kulturellen Themen aus. Ein deutlicher Unterschied zu großen Suchmaschinen ist die Antwortzeit: Selbst bei Lastspitzen liefern Google und Co. binnen drei Sekunden eine Ergebnisliste. Der YaCy-Anwender muss sich fünf bis fünfzehn Sekunden gedulden und die erzwungene Abfrage



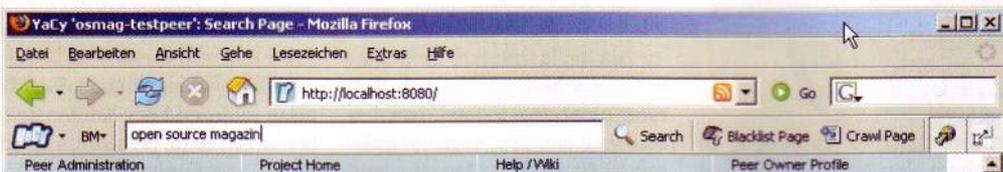
Benachbarte Peers zeigt eine automatisch erzeugte Grafik. Hier waren 62 Senior- und Principal-Peers online

fließt nicht nur der Besuch einer Seite in das Ranking ein, sondern auch die Meinung des Lesers nach dem Besuch. Wohl wichtigstes Kriterium für die Benutzung einer bestimmten Suchmaschine ist die – schwer messbare – Qualität der brauchbaren Ergebnisse. Bei Google sind wir allerhand Spam auf den ersten Ergebnisseiten gewöhnt, können aber immerhin eine große Anzahl Ergebnisse sichten und die für uns relevanten Informationen herausfiltern. Bei YaCy schwankt die Qualität enorm: Von lokalem Proxy (mit zusätzlicher Suchtiefe 1) bereits indexierte Seiten trugen bei einigen Themen zu sehr guten Suchergebnissen bei. Bei Topics, die traditionell viele Knotenbetreiber interessieren (Internet, Datenschutz, Open Source, Linux, Programmierthemen), ist die Er-

weiterer Peers kann noch einmal 30 Sekunden in Anspruch nehmen. Erst mit einer deutlich höheren Anzahl an Peers sind lange Ergebnislisten innerhalb von fünf Sekunden erreichbar.

Privatsphäre

Ein prinzipiell heikles Thema bei der Verwendung von Proxies als Suchmaschine ist der Schutz der Privatsphäre der Surfer. YaCy verwendet so genannte „Distributed Hash Tables“ – verteilte Indexe –, die in regelmäßigen Abständen mit anderen Hosts abgeglichen werden. So kann nicht zurückverfolgt werden, von welchem Host aus eine bestimmte Seite indexiert wurde. Und selbst wenn ein Eintrag ganz frisch ist, kann nicht zurückverfolgt werden, ob ein selbst gestarteter Crawl, ein Proxy-Abruf oder ein Remote-Crawl Ursache der Indexierung ist. Fast alle erwähnten Kritikpunkte sind auf die derzeit noch zu niedrige Knotenzahl zurückzuführen: Wer seine Lieblingsthemen besser indexiert sehen möchte, muss einen eigenen Knoten aufsetzen, der fortan sein Schärfelein beisteuert. Dies



Der YaCyBar für Firefox bietet neben einem Suchfeld schnellen Zugriff auf Funktionen wie „Social Bookmarking“ und „Quick Crawls“. Sie erhalten ihn unter www.yacy-websuche.de/wiki/index.php/De:YaCyBar

bringt uns zur Frage, wie groß das YaCy-Netz denn sein muss, um tatsächlich gegen Google und Co. bestehen zu können. Die YaCy-Entwickler gehen von einer vierstelligen Zahl an Senior-Peers im deutschsprachigen Raum und einer fünfstelligen Zahl weltweit aus. Erst dann sind einerseits große Teile des Webs indexiert und anderer-

seits genügend benachbarte Knoten vorhanden, um schnell Ergebnisse austauschen zu können. Als dieser Artikel entstand, waren in der Regel zwischen 50 und 100 Senior-Peers sichtbar – sollte YaCy jedoch in kurzer Zeit genauso populär werden wie viele Distributed-Computing-Projekte, kann diese Zahl in kürzester Zeit rapide ansteigen.

Fazit

YaCy ist ein interessantes Projekt, das derzeit weniger als Ersatz für die großen Suchmaschinen zu sehen ist, sondern als sinnvolle Ergänzung. Von Features wie „Social Bookmarking“ und der Proxy-Funktionalität profitieren insbesondere Netze, in denen sich mehrere Rechner

den Internetzugang über einen YaCy-Host teilen. Der Anwender profitiert vor allem von der wachsenden Konkurrenz: Nicht mehr Google legt fest, welche Kriterien das Ranking einer Seite bestimmen, sondern die Nutzerschaft, die ihre Rankingregeln zusammen mit den YaCy-Entwicklern selbst erarbeiten kann. ■

YaCy als Webserver

Wiki und Blog nutzen

Als Kind eines kollaborativen Verständnisses des Webs bietet YaCy ein Wiki, Blog-Funktionen und einen gewöhnlichen Webserver.

Etwas versteckt ist der gewöhnliche Webserver von YaCy. Um ihn zu aktivieren, genügt es, eine Datei „index.html“ im Ordner „DATA/HTDOCS“ des YaCy-Installationspfades abzulegen. Die eigene Webseite ist dann unter

```
http://ip-adresse:8080/www/
```

abrufbar. Etwas fixer einzurichten sind Blog und Wiki, beide zu finden unter „Kommunikation/Publikation“. Diese verwenden den Seitenaufbau und die Styles des YaCy-Layouts, erfordern also keine umfangreiche Anpassung. Änderungen an Blog und Wiki können über das YaCy-Newssystem (derzeit im Test) an andere Peers übermittelt werden. So ist auch im YaCy-Netz eine schnelle Verbreitung der eigenen Seiten sichergestellt.

YaCy als Webspider

Crawl starten

Alleine mit der Proxy-Funktionalität indexiert YaCy zwar bei den Anwendern bekannte und häufig nachgefragte Seiten, lernt aber nur wenig Neues dazu. Deshalb sind auch im YaCy-Netz gelegentliche Crawls sinnvoll.

Einen eigenen Crawl starten Sie unter „Index erzeugen“. Hierbei sind einige Kleinigkeiten zu beachten, denn in der Regel helfen andere Peers automatisch mit. Deren Administratoren sollten Sie nicht durch zu tiefe

oder zu häufige Crawls vergrätzen, sonst laufen Sie Gefahr, dass Ihre Anfragen für Crawls zurückgewiesen werden.

Zu beachten sind zwei Einstellungen: Zum einen die Crawl-Tiefe, die besagt, wie vielen Links der Crawl folgt, bis er stoppt. Da die meisten Seiten zehn bis 100 Links enthalten, ist mit einer Crawl-Tiefe von 5 bis 7 das gesamte Internet erreichbar, wenn die Suche nicht auf bestimmte Domains eingeschränkt wird. Hier kommt die Domain-Tiefe „Auto-Dom-Filter“ ins Spiel. Setzen Sie diesen Wert auf „1“, werden nur die Startdomain und unmittelbar von dieser Domain aus verlinkte Domains indexiert. Gute Erfahrungen haben wir übrigens mit der Kombination einer Crawl-Tiefe von 4 mit einer Domain-Tiefe von 1 gemacht. Startet man Crawls üblicherweise von Portalseiten aus, sollte die Crawl-Tiefe 3 nicht überschreiten.

YaCy als Adblocker und Zugangsschutz

Werbung blocken

Ein gut konfigurierter Proxy kann den Pop-up-Blocker sparen und hilft nebenbei, den gemeinsamen Index von Suchmaschinen-spamming – gegen das auch YaCy nicht gefeit ist – sauberzuhalten.

Um einen Überblick über Domains, die Werbebanner ausliefern, zu bekommen, sollten Sie nach einer ausgiebigen Surf-Session mit dem Link „Cache Monitor“ einen Blick in den Cache werfen. Verdächtig sind vor allem Domains, die Sie nicht explizit angesurft haben. Insbesondere unter „others“ finden Sie die Server der Anzeigenanbieter und können dort einen Blick in die von ihnen ausgelieferten Dateien werfen.

Unerwünschte Domains können Sie unter „URL Blacklist“ in eine eigene schwarze Liste eintragen. Bei den Blacklists sind Wildcards möglich, sodass Sie oft alle Server eines bestimmten Anbieters mit einem Muster wie

```
*.boeserwerber.*/*
```

erreichen. Blacklists stehen grundsätzlich allen anderen Peers offen, was wiederum bedeutet, dass auch Sie sich aus fremden Blacklists bedienen können. Leider zeigt das Drop-down-Menü nicht an, welcher Host eine nicht leere Liste vorhält.

YaCy als Webproxy

Speicher erhöhen

Bei unseren Experimenten mit YaCy haben wir festgestellt, dass das eingestellte Speicherlimit von 64 Megabyte oft zu knapp bemessen ist: Soll YaCy per „Prefetch“ verlinkte Seiten laden oder surfen mehrere User gleichzeitig über einen Proxy, muss deshalb der Speicher erhöht werden.

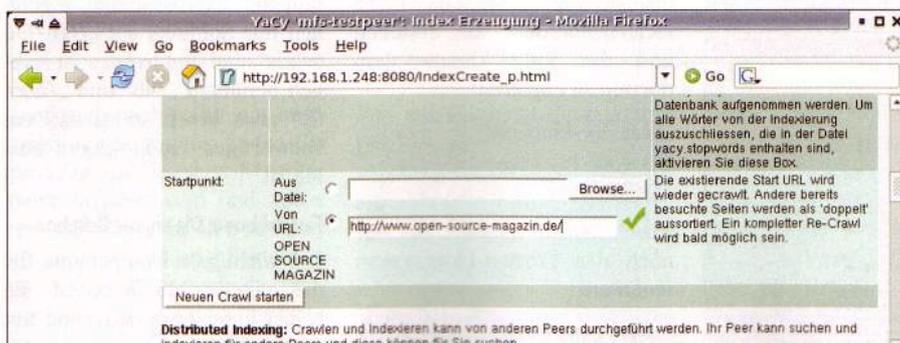
Wir haben hierfür die Datei „C:\Programme\YaCy\startYACY.bat“ im Editor geöffnet und die Zeile

```
set javacmd=-Xmx64m -Xms10m
```

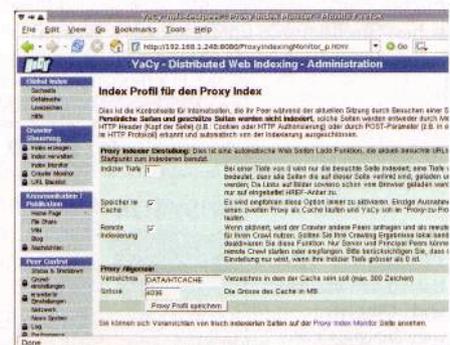
abgeändert auf:

```
set javacmd=-Xmx128m -Xms10m
```

Reicht die Verdoppelung noch nicht, erhöhen Sie den Parameter weiter. In einer unserer Testumgebungen erhielten wir erst ab 384 MB einen mehrere Tage am Stück problemlos arbeitenden YaCy-Peer.



Interessante Seiten, die nur unvollständig über den Proxy indexiert werden, können Sie crawlen lassen



Stellt man im Proxy eine Prefetch-Tiefe von 1 ein, indexiert YaCy nicht nur besuchte Seiten, sondern auch von diesen direkt verlinkte