

beispielsweise zum Lesen der Startseite 15 Minuten Zeit lasse, bevor ich auf eine weitere Seite (z. B. „Aktuell“) gehe, so werden hierdurch in der Server-Statistik zwei einzelne Besuche gezählt, obwohl es sich noch immer um die gleiche Sitzung handelt. Insofern kann also der Messwert der *visits* keine exakte Angabe einzelner Besuche sein, kommt diesem Wert jedoch in der Regel recht nahe.

Aus dem hier gesagten geht also auch hervor, dass die Angabe der einzelnen (verschiedenen) Rechner nicht exakt sein kann, da sie nur zählt, wie viele verschiedene IP-Adressen die Website besucht haben. Wenn ich mich von zu Hause aus an einem Tag zweimal ins Netz einwähle, werde ich mit großer Wahrscheinlichkeit bei beiden Sitzungen zwei verschiedene IP-Adressen zugewiesen bekommen und so

mit erhöhe ich hiermit auch die Anzahl einzelner Rechner in der Server-Statistik um zwei, obwohl ich lediglich zweimal mit dem gleichen Rechner die Website besucht habe. Andersherum geht es natürlich auch: Wenn sich fünf verschiedene Angestellte unseres Büros von ihren jeweiligen Arbeitsplatz-Rechnern aus unsere Website ansehen, so sind dies tatsächlich fünf verschiedene Rechner, die jedoch in der Statistik lediglich als ein Rechner gezählt werden. Finden diese Besuche auch noch gleichzeitig statt, so wird auch nur ein *visit* gezählt, obwohl es sich um fünf verschiedene *visits* handelt.

So, nun seht euch mal die erste Seite der Server-Statistik an und macht euch mal selbst ein Bild:

„<http://www.speleo-berlin.de/stats>“

Roman Blöth

Letzte Meldung

Ein Telefonat mit K. H. Pielsticker (Speläogruppe Letmathe) über eine kürzlich durchgeführte Expe nach Ägypten brachte für meine Begriffe absolut sensationelle Neuigkeiten ans Licht: Das mit Geologen und Höhlenkundlern hochkarätig besetzte Team fand mitten in der Wüste frei in der Landschaft stehende Stalagmiten. Teilweise handelte es sich um „Palmenstamm“-Stalagmiten vom Typ „Aven Organac“ (Südfrankreich), die nur entstehen können, wenn große Tropfenfallhöhen vorkommen.

Von Höhlen ist ringsherum nichts mehr zu sehen. Die Erklärung für dieses weltweit einmalige Phänomen (zudem noch in einem Land, wo heute fast keine bedeutenden Höhlen bekannt sind): Durch Deckeneinbrüche ehemaliger großer Karsthölräume mit reichem Tropfsteinschmuck

drang Sand ein und erfüllte die Höhlen völlig. Anschließend wurde durch die Erosion (ca. ein Millimeter pro Jahr sind anzunehmen) die gesamte Höhle abgetragen, die Tropfsteine blieben aber von der Sandfüllung geschützt. Datierungen der Tropfsteine müssen indirekt über Sedimentdatierungen erfolgen, da wegen dem extremen Alter (es dürfte sich um die ältesten Tropfsteine der Welt handeln) selbst die Uran/Thorium-Methode versagt.

Michael Laumanns

Impressum

SCB-Newsletter, Nr. 7, November 2001, 1. Jahrgang
ISSN 1618-4785, www.speleo-berlin.de
unregelmäßig erscheinendes Nachrichtenblättchen des Speläoclub Berlin,
c/o Torsten Kohn, Konrad-Wolf-Straße 13b,
13055 Berlin, tkohn@speleo-berlin.de
Redaktion: Norbert Marwan
Kastanienallee 17, 14471 Potsdam,
marwan@speleo-berlin.de

Editorial

Der November ist noch gar nicht um, der sechste Newsletter noch nicht mal veröffentlicht, und schon wieder ist ein SCBNL gefüllt. Neben Kurzberichten vom Harz und aus Ägypten gibt es die Fortsetzung zum Thema Server-Statistik.

Die Schar von Persien...

... traf sich am Wochenende vom 19. bis 21. Oktober zu einer Nachbesprechung im Vereinsheim der Harzer Höhlenforscher in Rübeland. Dort fanden sich zahlreiche Harzer Höfen und von den Iranfahrern Bernhard Köppen, Georg Kaufmann, Ingo Dorsten und der Verfasser ein. Zusätzlich stieß am Samstag früh noch der Deutschrusse Nikolai mit Familie aus Gießen dazu, der das Treffen durch seine angenehme Art und sein großes Fachwissen zusätzlich bereicherte.

Der SCB im WWW

<http://www.speleo-berlin.de>

Nach der Beisetzung von André Hörchner am 19. Oktober in Suhl, die mich – auch wenn ich persönlich nicht eng mit André befreundet war – wie alle anderen Anwesenden auch sehr bewegt hat, war ich wirklich froh, am gleichen Abend noch auf andere Gedanken zu kommen, sonst wäre es trotz des herrlichen Herbstwetters für mich noch ein ganz depressives Wochenende geworden. Lasst uns hoffen, dass ähnliche schreckliche Unfälle in Zukunft nicht mehr passieren!

Die Harzer Kollegen – allen voran Ingo Dorsten, Siegfried Wielert und Familie Schuhose – hatten am Samstag ein tolles Exkursionsprogramm zusammengestellt, das wir bei phantastischem Wetter rings um Rübeland zu Fuß absolvierten. Am Vormittag ging es im Rahmen einer Rundwanderung rings um den Rübeländer Karst, anschließend durften wir kostenlos den Führungsteil der Hermannshöhle besichtigen und am späten Nachmittag gab's dann noch ein echtes Highlight: Die Kameruner Höhle. Nach einem relativ kurzen engen Eingangsteil, der in den 50er Jahren von Rübeländer Speläologen aufgegraben wurde, gelangt man in sehr geräumige hallenartige Gänge mit außerordentlichem Tropfsteinschmuck, der in einer unerschlossenen deutschen Höhle wirklich bemerkenswert gut erhalten ist. Sowa sieht man wirklich nicht alle Tage! Einer der großen Gänge hat mehrere Basaltintrusionen „durchschlagen“, was es in dieser Form auch nirgends in Deutschland zu sehen gibt.

Der Samstagabend wurde mit Iranbildern und Fachsimpelei über den Höhlenplan der Ghar Alisadr gefüllt, die nunmehr eine offizielle Länge von 11.440 m hat.

Mit diesem Wochenende hat sich Rübeland für mich als Veranstaltungsort für das 75jährige Jubiläum der Harzer Höhlenforschung – eine große Veranstaltung ist zu Pfingsten 2003 geplant – nachhaltig empfohlen. Da gibt es noch viel zu sehen und ich bedanke mich herzlich bei allen Harzer Kollegen für das tolle Wochenende.

Michael Laumanns

Fortbildung im SCB

Server-Statistik lesen und verstehen II

Nachdem wir im ersten Teil den Begriff der *hits* beschrieben haben, wird nun im zweiten Teil die Maßeinheit *page impressions* (bzw. *page views*, was das Gleiche ist) erklärt. Eine *page impression* wird immer dann gemessen, wenn eine einzelne HTML-Datei an einen Besucher der Website übertragen wurde, unabhängig davon, wie viele Bilder oder sonstige zusätzliche Dateien zum „Inhalt“ der betreffenden HTML-Datei gehören. Die Messung der *page impressions* sagt also etwas darüber aus, wie viele Seiten auf einer Website in einem bestimmten Zeitraum (wahrscheinlich) von Web-Besuchern angesehen wurden. Natürlich kann hiermit nicht gemessen werden, ob diese Seiten auch gelesen wurden, aber selbst wenn man das könnte, würde dies noch nichts darüber aussagen, ob die gelesenen Seiten auch verstanden wurden. Insofern ist die Maßeinheit *page impressions* also die grundlegend wichtigste zur Interpretation der Server-Statistik. Mit diesem Wissen können wir nun bereits die Startseite unserer Server-Statistik verstehen.

In der unteren Tabelle werden die wichtigsten Angaben – nach Monat aufgeschlüsselt – auf einen Blick sichtbar. Die linke Hälfte der Tabelle zeigt die jeweils über den betreffenden Monat gemittelten Tagesdurchschnittswerte, die rechte Hälfte zeigt die Werte als monatliche Summen an.

Beginnen wir mit den einzelnen Spalten von links nach rechts:

1. Tagesdurchschnitt/ Anfragen. Hier werden die durchschnittlichen *requests* (englisch: Anfragen, Anforderungen) aufgelistet. Soll mein Browser eine HTML-Seite mit vier Bildern anzeigen, so wird er an den Webserver insgesamt fünf Anfragen

stellen. Ob diese auch mit den jeweils angefragten Dateien beantwortet wurden, sieht man in der nächsten Spalte:

2. Tagesdurchschnitt/ Dateien. Diese Angabe entspricht den tatsächlichen *hits*, also der Anzahl tatsächlich übermittelter Dateien. Wenn ich mir eine Seite regelmäßig ansehe, so werden sich die betreffenden Dateien aller Wahrscheinlichkeit nach bereits im Cache meines Browsers befinden, so dass sich dieser diese Dateien nicht mehr vom Server übermittelt lassen muss, sondern sich gleich aus dem Cache holen kann. Aus diesem Grunde ist die Anzahl der übermittelten Dateien geringer als die Anzahl der angefragten Dateien. Ob der Browser nämlich eine Datei aus seinem Cache holen kann oder vom Server (erneut) anfordern muss, kann er erst dann entscheiden, wenn der Server die betreffende Anfrage beantwortet hat (aus dieser Antwort geht nämlich hervor, ob die angefragte Datei auf dem Server mit der im Cache noch identisch ist oder inzwischen verändert wurde).

3. Tagesdurchschnitt/ Seiten. Diese Spalte gibt die oben beschriebenen *page views* an und ist somit der interessantere aller bisher genannten Werte.

4. Tagesdurchschnitt/ Besuche. Diese Spalte soll angeben, wie viele einzelne *visits* (englisch: Besuche) es (durchschnittlich pro Tag) auf der Website gab. Dieser Wert ist allerdings in der Regel nicht exakt bestimmbar. Hierzu folgt im Weiteren eine genauere Beschreibung.

5. Tagesdurchschnitt/ Rechner. Misst, wie viele verschiedene Rechner (gemeint sind die verschiedenen IP-Adressen der Besucher der Website) die Website durchschnittlich pro Tag besucht haben. Auch dieser Wert ist nicht 100%ig genau messbar.

6. Tagesdurchschnitt/ kb. Bedarf wohl kaum einer Erklärung und ist lediglich interessant, um den sogenannten *traffic* auf der Website, also die tatsächlich vom Server aus übertragene Datenmenge, zu bestimmen. Glücklicherweise haben wir in unserer Web-Präsenz (also bei unserem Internet Service Provider, welcher uns den Rechnerplatz und dessen Internet-Anbindung zur Verfügung stellt) ein recht hohes monatliches Datentransfer-Limit. Die eigentlich im Internet entstehenden Kosten richten sich nämlich nach dem Datenverkehr, also der übertragenen Datenmenge, vom betreffenden Server aus. Die Service Provider zahlen Geld pro monatlich übertragene Datenmenge.

Die hohen Werte im ersten Monat unserer neuen Web-Präsenz (Mai 2001) ergeben sich aus der einfachen Tatsache, dass ich in dieser Zeit sehr viel auf die Website zugegriffen habe, um die Seiten zu erstellen und zu aktualisieren. Nehmen wir diese Zeile also einmal aus der Betrachtung heraus, so sehen wir, dass die Anzahl der Zugriffe/ Besuche auf unserer Website in den letzten Monaten recht gut zugenommen haben – durchschnittlich 9 Besucher pro Tag im Oktober finde ich für eine so spezielle Website wie unsere doch schon recht beachtlich: Unsere Seiten werden also tatsächlich gelesen!

Zu den „gemessenen“ Besuchen (*visits*) muss ich zunächst einmal erläutern, wie diese Angaben überhaupt zustande kommen. Als Otto-Normalverbraucher wähle ich mich mit einem Modem oder einer ISDN-Karte jedesmal ins Internet ein, wenn ich mir Websites ansehen möchte. Grundsätzlich benötigt jeder mit dem Internet verbundene Rechner eine sogenannte IP-Adresse, die weltweit einmalig nur an diesen Rechner vergeben ist. Anbieter wir z. B. T-Online verfügen hierzu über ei-

nen größeren Pool an IP-Nummern, die dynamisch der Reihe nach an die sich bei T-Online einwählenden Kunden vergeben werden. Wird nach Beendigung einer Sitzung die Verbindung zum Internet unterbrochen, so wird die dabei verwendete IP-Adresse wieder frei und kann an den nächsten Kunden, der sich einwählt, vergeben werden. Hat ein Rechner hingegen eine Standleitung, so wird sich seine IP-Adresse in der Regel langfristig nicht ändern und steht somit für einen bestimmten Benutzer oder wenigstens eine bestimmte Benutzergruppe. In vernetzten Büros dient in der Regel ein Zugangsrechner als Schnittstelle zwischen den einzelnen Arbeitsplatzrechnern und dem Internet. Aus Sicht „von Außen“ benutzen hierbei alle einzelnen Rechner aus diesem Büro die gleiche IP-Adresse.

Die Annahme bei der Erzeugung der Server-Statistiken ist nun die, dass ein einzelner Besuch dadurch gekennzeichnet ist, dass von einer bestimmten IP-Adresse aus im Abstand von maximal z. B. fünf Minuten einzelne *hits* bzw. *requests* generiert werden. Fordert eine IP-Adresse z. B. um 14:00 Uhr die Startseite *index.html* und im Abstand von Millisekunden noch die dort eingebundenen Bilder an, und nach ca. einer Minute die Dateien einer weiteren HTML-Seite, so wird angenommen, dass es sich hierbei noch um den gleichen Besucher handelt. Ist jedoch eine bestimmte Zeitspanne vergangen, nach der wieder von der gleichen IP-Adresse Dateien vom Server angefordert werden, so nimmt man an, dass es sich hierbei um einen neuen Besuch der Website handelt, wobei aufgrund der zuvor genannten Umstände nicht gesagt werden kann, ob es sich um einen neuen Besucher handelt, oder um die gleiche Person wie zuvor, die sich lediglich erneut die Website ansieht. Wenn ich mir nun