

ホームページサービスの試行(2)

伊藤 彰朗*

1. はじめに

前回の広報(ホームページサービスの試行, Vol.3 4, No.2)では、ホームページサービスの概要とサービスの内容について説明しました。今回は、本サービスで提供するアクセスカウンター及びログ参照機能について説明します。

2. アクセスカウンター

本サービスで提供するアクセスカウンターは、

図1で示すようなグラフィック文字によるカウンターです。利用するには、次のような、1行をWebページに記入します。

あなたは番目の利用者です。

グラフィック文字の大きさは、高さ指定のheightのみで行います。本稿では、使用例が見やすいようにグラフィック文字を大きく(height=30)しましたが、heightは12~18ぐらいが適当かと思います。

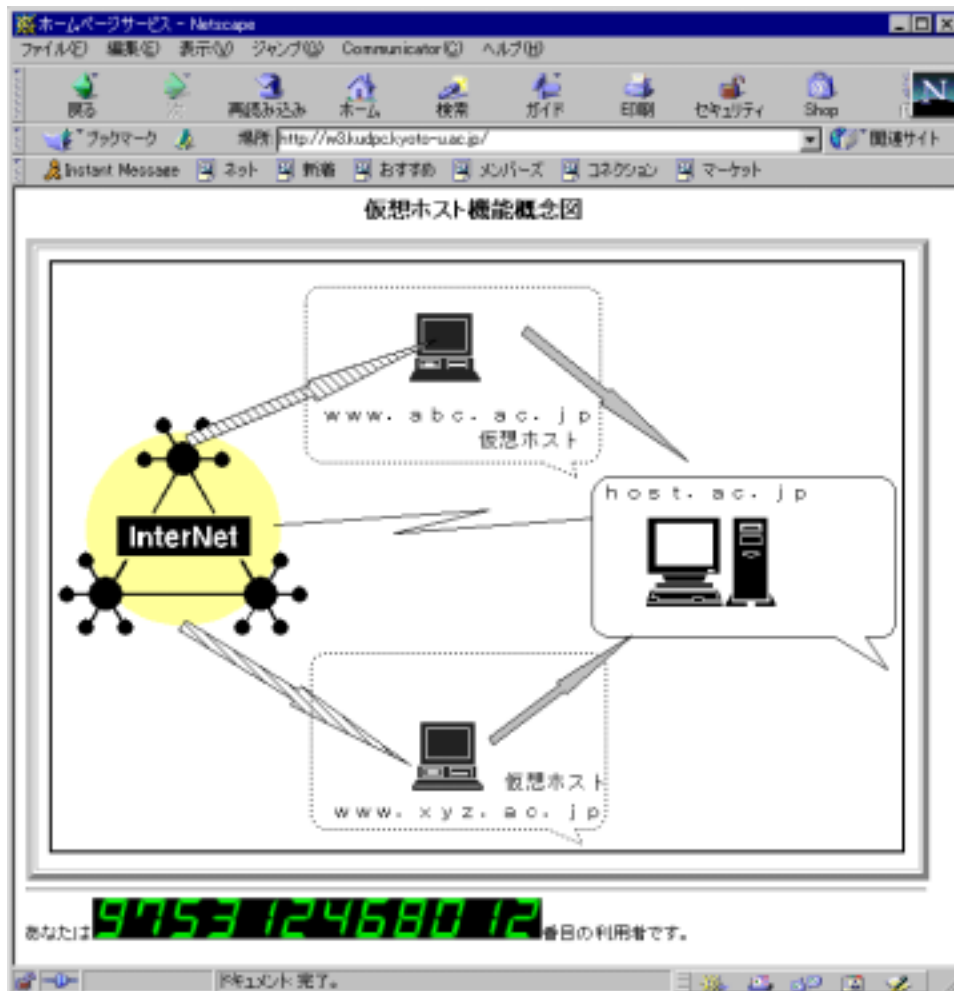


図1 アクセスカウンター使用例

*いとう あきろう(京都大学大型計算機センター)

3. ログ参照機能

本サービスで提供するログ参照機能では、Free softwareであるアクセスログ解析プログラムのwwwstatとanalogを利用し、アクセスログの解析結果を提供をします。ここでは本センターの1ヶ月分データをいどういった分析が可能か例示します。

3.1 wwwstatによるアクセスログの解析

wwwstatによる解析は、WWWサーバがアクセス毎に保存しているアクセスログを次に示す6種類のカテゴリで集計し、HTML文章として作成することができます。図2～図8にwwwstatによるアクセスログの解析結果を表示します。本サービスの申請者へのwwwstatによる解析データについては、1ヶ月毎にHTML文章ファイルで提供します。なお、wwwstatの解析には、いくつかのオプションがあります。希望されるオプションは、ご相談に応じます。

Totals for Summary Period

図2に解析期間内（この例では1ヶ月とした）で集約し、アクセス状況を以下の4項目で示します。

- 全ファイル転送数
- 全データ転送量
- 平均ファイル転送数
- 平均データ転送量

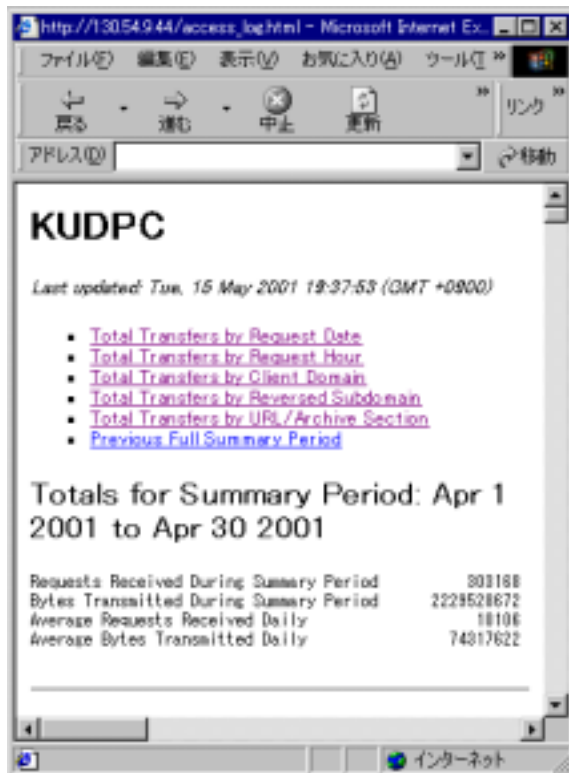


図2 1ヶ月のアクセス状況

Total Transfers by Request Date

図3に日毎のアクセス状況を以下の5項目で示します。

- 当日リクエスト数 ÷ 全リクエスト数（期間内）
- 当日転送データ量 ÷ 全転送データ量（期間内）
- 当日転送データ量
- 当日リクエスト数
- 日付



図3 日毎のアクセス状況

Total Transfers by Request Hour

図4に時間帯毎のアクセス状況を示します。なお、表示するためのソート項目は3種類選択できます。デフォルト（特に指定のない場合は）、右端の項目（時刻）でソートされますが、オプションでリクエスト数、データ量でのソートも可能です。

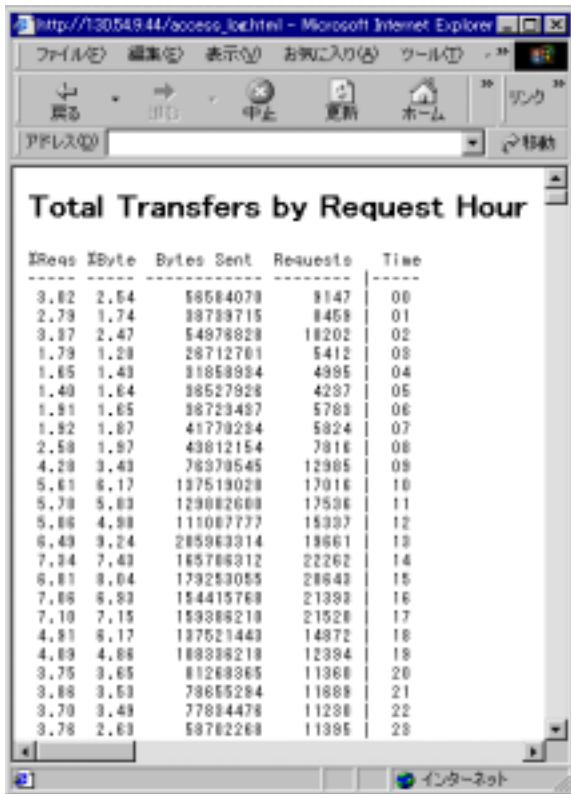


図4 時間帯毎のアクセス状況

Total Transfers by Client Domain

図5にドメインによる国別のアクセス状況を示します。なお、表示するためのソート項目は3種類選択できます。デフォルトは、右端の項目(国別)でソートされますが、リクエスト数、データ量でのソートも可能です。



図5 国別アクセス状況

Total Transfers by Reversed Subdomain

図6にサブドメイン別のアクセス状況を示します。IPアドレスがDNSにあればホスト名で表示され、DNSにない場合はIPアドレスで表示されます。wwwstatではホスト名別の集計(図7)と図6のようにホスト名からマシン名を取り除いてサブドメインとした集計の選択ができます。図7のホスト名による集計では、矢印で示したように詳細に集計されるが、サブドメイン別とすれば図6の罫線枠で示したように簡潔に集計できます。なお、京大のドメインとそれ以外のドメインに区別し、ホスト名集計またはサブドメイン集計のどちらかの選択もオプションでできます。図6の処理がデフォルトです。ドメイン名などの表記は、DNSによりホスト名を参照するため長い処理時間を必要とします。このサービスは、次に説明するanalogでは行いません。

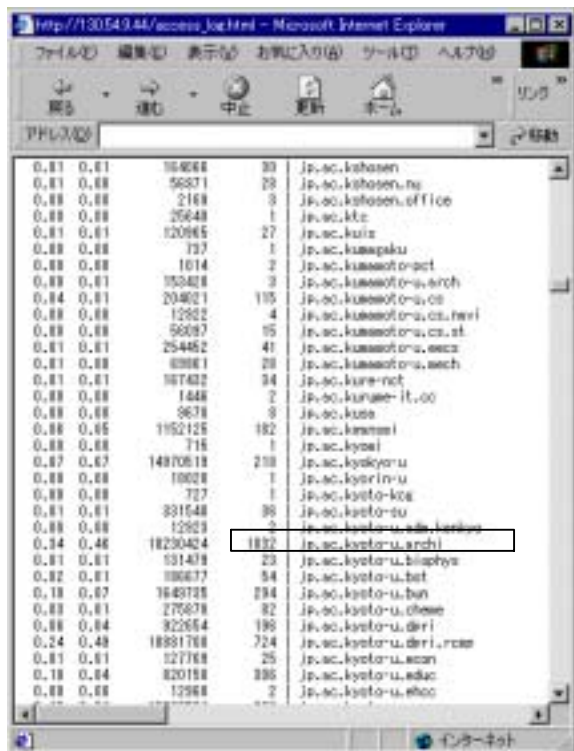


図6 サブドメイン別のアクセス状況

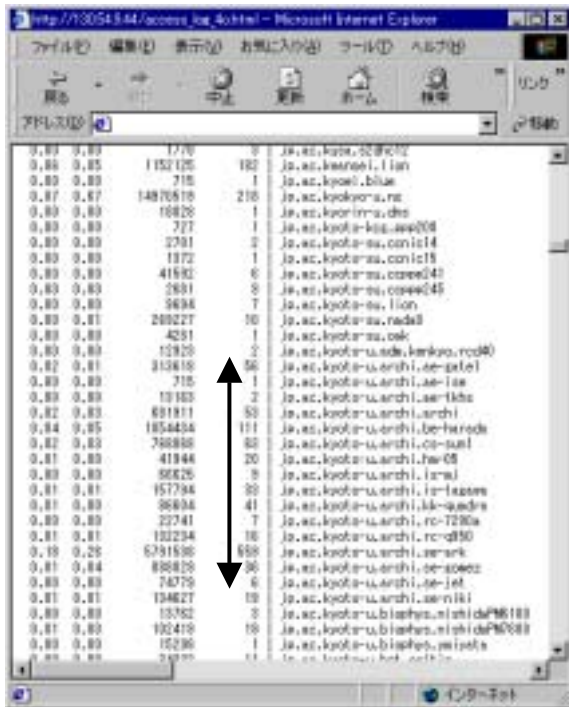


図7 ホスト名別のアクセス状況

Total Transfers by URL/Archive Section

図8にファイル毎のアクセス状況を示します。オプションでファイルへのハイパーリンクをつけることができます。

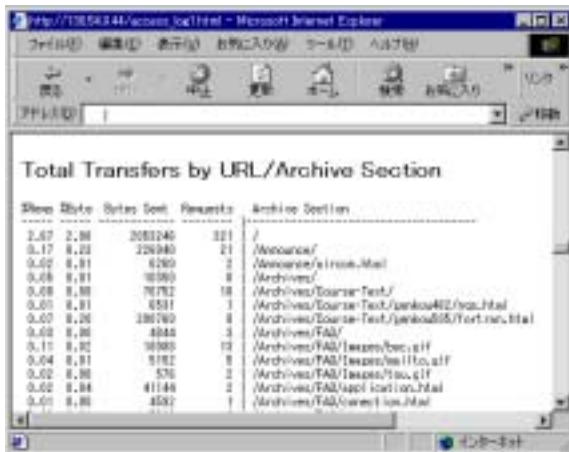


図8 ファイル毎のアクセス状況

3.2 analogによるアクセスログの解析

analogによる解析は、ブラウザー画面による選択によりアクセスログを集計し、図10～図19のようなグラフィカルなアクセス状況をリアルタイムに表示できます。wwwstatの図6のようにサブドメイン名での表記は処理時間がかかることとサーバ負担の関係で行えません。

選択画面

図9のような16種類の選択で集計できます。



図9 選択画面

全体の概要集計

図10に期間内の集計をリクエスト件数など12種類の項目で示します。なお、画像ファイル(jpg, gif, bmp)に関しては、リクエスト数に反映されないように解析から除外しているためwwwstatで集計した件数とは、その分少なくなります。analogで使われる「リクエスト」は、外部からのアクセス回数を意味し、「ページ」は、HTMLページかディレクトリが表示されたときのみ1ページです。



図10 全体の概要集計

月別、週別、曜日別集計

月別、週別を図 1 1 に示します。期間内のリクエスト数を棒グラフで示します。なお、最も混雑した週が下段に表示され、4月22日の週が最も混雑していることがわかります。図 1 2 の曜日別では、休日のアクセスが、平日のピークより30%以上少ないことがわかります。



図 1 1 月別、週別

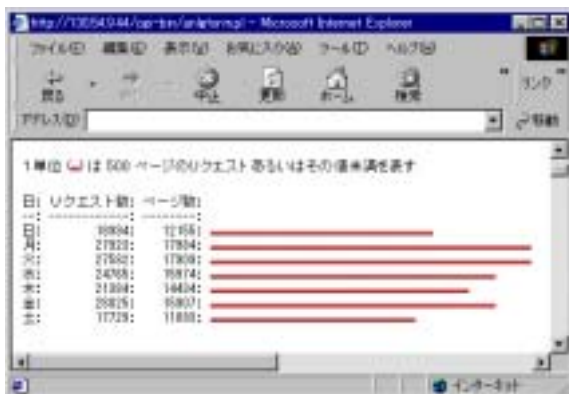


図 1 2 曜日別

日別、時間帯別集計

日別を図 1 3 に示します。4月24日に最も多くのページが参照されたことがわかります。時間帯別を図 1 4 に示します。14時の時間帯が最も多くのページが参照されています。



図 1 3 日別



図 1 4 時間帯別

IPアドレス別

図15にIPアドレス別を円グラフで表示します。
図16のようにIPアドレスを階層的に表示することができます。



図15 IPアドレス別(1)

ファイルサイズ別

図17にファイルサイズ別を示します。ファイルサイズで1kb~10kbまでのリクエスト件数が50%近くあることが円グラフからわかります。



図17 ファイルサイズ別

ファイル種類別

図18にリクエストされた拡張子のファイル種類別を示します。通信量とはデータ転送量でバイト数です。全体の割合(%)で各ファイル種類別通信量を円グラフに表示します。画像データは除外しているので含まれないが、pdf、htmlで80%近くの通信量を占めています。

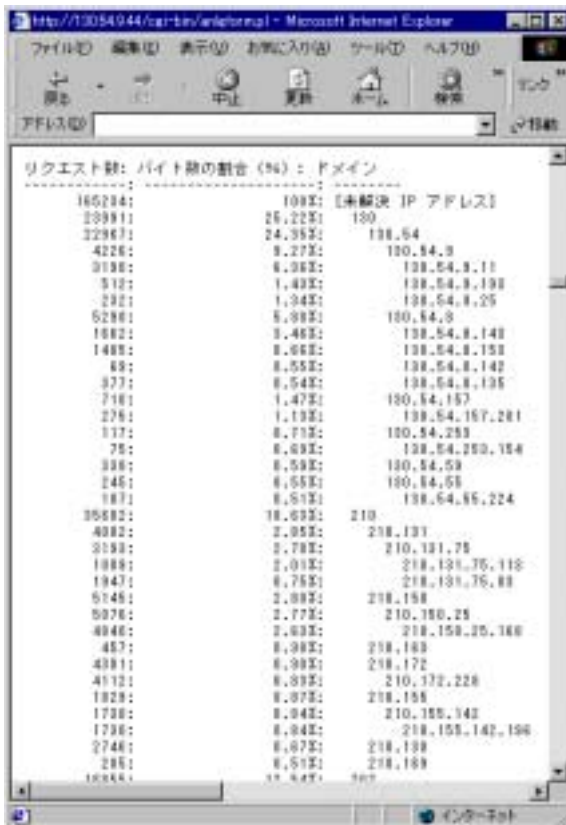


図16 IPアドレス別(2)

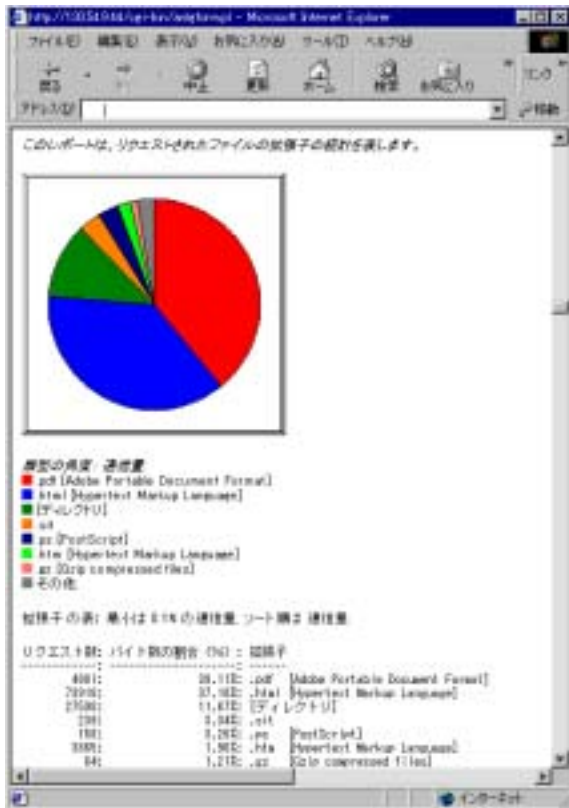


図 18 ファイル種別別

ディレクトリ別

図 19 にディレクトリ別を示します。/Services/ というディレクトリが全体の通信量の 40% あることが円グラフからわかります。/Fujitsu_Manual/ も 16% あることから、Web から本センターのサービス、計算機のマニュアルに関する情報を入手する目的であることが伺えます。



図 19 ディレクトリ別

4. おわりに

本稿では、本サービスで提供するアクセスカウンターの使い方、ログ参照機能については、分析用途も含めて具体例を挙げた説明をしました。ログ参照機能についての使い方やオプションによる解析データの選択など、具体的なことは申請者にご相談して提供できるように考えています。

