***NASA-HTTP***

**Description**

These two traces contain two month's worth of all HTTP requests to the NASA Kennedy Space Center WWW server in Florida.

**Format**

The logs are an ASCII file with one line per request, with the following columns:

1. **host** making the request. A hostname when possible, otherwise the Internet address if the name could not be looked up.
2. **timestamp** in the format "DAY MON DD HH:MM:SS YYYY", where **DAY** is the day of the week, **MON** is the name of the month, **DD** is the day of the month, **HH:MM:SS** is the time of day using a 24-hour clock, and **YYYY** is the year. The timezone is -0400.
3. **request** given in quotes.
4. **HTTP reply code**.
5. **bytes in the reply**.

**Measurement**

The first log was collected from 00:00:00 July 1, 1995 through 23:59:59 July 31, 1995, a total of 31 days. The second log was collected from 00:00:00 August 1, 1995 through 23:59:59 Agust 31, 1995, a total of 7 days. In this two week period there were 3,461,612 requests. Timestamps have 1 second resolution. Note that from 01/Aug/1995:14:52:01 until 03/Aug/1995:04:36:13 there are no accesses recorded, as the Web server was shut down, due to Hurricane Erin.

**Privacy**

The logs fully preserve the originating host and HTTP request. Please do not however attempt any analysis beyond general traffic patterns.

**Acknowledgements**

The logs was collected by Jim Dumoulin of the Kennedy Space Center, and contributed by Martin Arlitt (*mfa126@cs.usask.ca*) and Carey Williamson (*carey@cs.usask.ca*) of the University of Saskatchewan.

**Publications**

This is one of six data sets analyzed in an upcoming paper by   
M. Arlitt and C. Williamson, entitled ``Web Server Workload Characterization: The Search for Invariants'', to appear in the proceedings of the *1996 ACM SIGMETRICS Conference on the Measurement and Modeling of Computer Systems*, Philadelphia, PA, May 23-26, 1996. An [*extended version*](ftp://ftp.cs.usask.ca/pub/discus/paper.96-3.ps.Z) of this paper is available on-line; see also the [*DISCUS home page*](http://www.cs.usask.ca/projects/discus/discus.html) and the group's [*publications*](http://www.cs.usask.ca/projects/discus/discus_pub.html).

**Related**

Permission has been granted to make four of the six data sets discussed in ``Web Server Workload Characterization: The Search for Invariants'' available. The four data sets are: [Calgary-HTTP](http://ita.ee.lbl.gov/html/contrib/Calgary-HTTP.html), [ClarkNet-HTTP](http://ita.ee.lbl.gov/html/contrib/ClarkNet-HTTP.html), [NASA-HTTP](http://ita.ee.lbl.gov/html/contrib/NASA-HTTP.html), and [Saskatchewan-HTTP](http://ita.ee.lbl.gov/html/contrib/Sask-HTTP.html).

**Restrictions**

The traces may be freely redistributed.

**Distribution**

Available from the Archive in [**Jul 01 to Jul 31, ASCII format, 20.7 MB gzip compressed**](ftp://ita.ee.lbl.gov/traces/NASA_access_log_Jul95.gz), 205.2 MB uncompressed, and [**Aug 04 to Aug 31, ASCII format, 21.8 MB gzip compressed**](ftp://ita.ee.lbl.gov/traces/NASA_access_log_Aug95.gz), 167.8 MB uncompressed.

資料來源 http://ita.ee.lbl.gov/html/contrib/NASA-HTTP.html