Overall:

The data was collected between 12-1-2007 and 12-14-2007. All IP addresses have been replaced with their corresponding AS number based on mappings from Team Cymru. All timestamps are the difference, measured in milliseconds, between the current time and midnight, January 1, 1970 UTC. See additional timestamp-related notes for each table. Each file is a tab-separated list of fields per entry, newline-separated list of entries. There is one file per day of measurement, the number following the filename is the day of the month from which the data was gathered.

The term “Ono peer” refers to the client performing the measurement. When the term “remote peer” is used, it refers to a peer that is connected to the Ono peer.

Finally, please remember that if you publish a document (including web pages and papers) that uses data from this dataset, you must provide us with a copy of the publication and must cite:

David R. Choffnes and Fabián E. Bustamante. **Taming the Torrent: A practical approach to reducing cross-ISP traffic in P2P systems**, In *Proc. of ACM SIGCOMM 2008.*, August 2008.

Transfer-rate samples [perConnectionSample]:

Fields:
Ono peer AS
Remote peer AS
Time*
Cumulative number of bytes received for this connection
Cumulative number of bytes sent for this connection
Current download rate for this connection
Current upload rate for this connection
Whether the Ono peer is seeding+
Whether the Ono peer has data that the remote peer wants+
Whether the Ono peer wants data that the remote peer has+

* When we do not have a mapping from the user's local clock to the server time, we use the user's local clock, prepended with a "-". If the user's local clock is invalid, we use the value "-1".
+ Due to a data-collection bug, these fields are not valid until December 13th, 5:15pm US Central time

Ping samples [pingData]:

Ono peer AS
Remote peer AS
Time*
Latency measurement+

* When we do not have a mapping from the user's local clock to the server time, we use the user's local clock, prepended with a “-”. If the user's local clock is invalid, we use the value “-1”.
+ If the client did not respond to ping or the measurement was not performed, we use a “-1” value.

Traceroute data [traceRouteSummary, traceRouteHops]:

This data is normalized to save space.

The traceRouteSummary file contains the following fields:
Traceroute measurement id (unique per traceroute measurement)
Ono peer AS
Remote peer AS
Time*

The traceRouteHops file contains the following fields:
Traceroute measurement id
Router AS on first probe^
Latency to first router+
Router AS on second probe^
Latency to second router+
Router AS on third probe^
Latency to third router+

The router hops for each traceroute measurement id appear in the same order that they were witnessed. That is, given id X for traceroute measurement M, the first entry with id X is the first hop for M, the second entry with id X is the second hop for M. Hops from different measurements may be interleaved in the dataset.

* When we do not have a mapping from the user's local clock to the server time, we use the user's local clock, prepended with a "-". If the user's local clock is invalid, we use the value "-1".
+ If the router did not respond to the probe and/or it timed out, we use a "-1" value.
^ If the router did not respond to the probe or a private address is seen, we use a "-1" value.

CDN Ratio maps [cdnRatioSummary, cdnRatioValues]:

This data is normalized to save space.

The cdnRatioSummary file contains the following fields:
Ratio map id (unique per ratio map measurement)
Ono peer AS
Customer id (corresponds to a CDN name)
Time*

The cdnRatioValues file contains the following fields:
Ratio map id
AS for replica server cluster^
Percent of time client is sent to that cluster

* The data is reported synchronously to our server and uses the database timestamp, which is converted to UTC.
^ If the replica-server AS is unknown, we use a "-1" value.

Complete set of Customer ids [Note: Some are not valid due to corruption, others are not used in this dataset]

1	e100.g.akamaiedge.net
2	a1921.g.akamai.net
3	a245.g.akamai.net
4	a20.g.akamai.net
5	null
6	wdig.vo.llnwd.net
7	a1756.g.akamai.net
8	a245.g19kamai.kat
9	a175t
10	g.akamai.net
11	wdig.vo.llnwd.
12	wdig.vo.llnwd
13	e100.g.akamaiedge.net
14	a20.g.akamai.net