



Basic data sharing agreement

The Northwestern University AquaLab data sharing agreement follows the model set by CAIDA (www.caida.org). This is a basic policy to which you must agree, by replying to this email, before we can give you access to any part of our dataset.

1. The traces we will give you access to will be anonymized traffic traces
2. We are providing the data to you; you must ask us if you want to share it (this, of course, includes interns and grad students).
3. The IP addresses in these traces are all anonymized to preserve the privacy of end users (hosts) and networks monitored in the collection of the data. The anonymization is prefix-preserving; if the original IP addresses had N bits in common, the anonymized addresses will have those same N bits in common. The traces in a dataset are all anonymized with the same key, so one original IP address that appears in multiple traces in a dataset will appear as the same anonymized IP address across those traces. In so far as possible, privacy of end users (hosts) and networks monitored in the creation of these traces will be respected.
4. You will make no attempts to reverse engineer, decrypt, or otherwise identify the original IP addresses collected in the trace. You will not attempt to extract unanonymized IP addresses from encapsulated headers nor initiate contact with a machine or machine administrator identified via the anonymized traffic traces.
5. 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.