ip packet lengths with HDFS operations

Louis-Claude Canon
April 18, 2012

Traces captured with tcpdump -i eth0 -s 65535 -w /tmp/node.2.large.1.data (some packet were not captured) when writing 64 MiB to a single HDFS datanode.

> T <- read.table("node.2.large.1.data.gz", header = TRUE, sep = ",")
> plot(T[, "Time"], T[, "Length"], cex = 0.3)

The length of each IP packet increases with time (the gap may be related to the HDFS acknowledgement layer).

> index_first <- T[, "Time"] < 2.6
> index_last <- T[, "Time"] > 3.2
The ECDF for the first and last packets confirms the previous trend. When the file is read, all the packets have the same length: 1514 B.