Their share: diversity and disparity in IP traffic

Andre Broido[1], Young Hyun[1], Ruomei Gao[2], and kc claffy[1]

[1] Cooperative Association for Internet Data Analysis
SDSC, University of California, San Diego
{broido, youngh, kc}@caida.org
[2] Georgia Institute of Technology
gaorm@cc.gatech.edu

Abstract. The need to service populations of high diversity in the face of high disparity affects all aspects of network operation: planning, routing, engineering, security, and accounting. We analyze diversity/disparity from the perspective of selecting a boundary between mice and elephants in IP traffic aggregated by route, e.g., destination AS. Our goal is to find a concise quantifier of size disparity for IP addresses, prefixes, policy atoms and ASes, similar to the oft-quoted 80/20 split (e.g., 80% of volume in 20% of sources). We define *crossover* as the fraction c of total volume contributed by a complementary fraction $1 - c$ of large objects. Studying sources and sinks at two Tier 1 backbones and one university, we find that splits of 90/10 and 95/5 are common for IP traffic. We compare the crossover diversity to common analytic models for size distributions such as Pareto/Zipf. We find that AS traffic volumes (by byte) are top-heavy and can only be approximated by Pareto with $\alpha = 0.5$, and that empirical distributions are often close to Weibull with shape parameter 0.2–0.3. We also find that less than 20 ASes send or receive 50% of all traffic in both backbones' samples, a disparity that can simplify traffic engineering. Our results are useful for developers of traffic models, generators and simulators, for router testers and operators of high-speed networks.[3]

1 Introduction.

The lack of a predictive relation between *number* (cardinality) and the combined *size* (volume) of a collection of objects is a recurrent problem in Internet data analysis. The volume of a natural category, such as users, instants, and networks, can be unevenly distributed among individual objects. For example, while the Internet architecture does not have any *single* point of failure, 80-90% of routes use the top 20 providers. Similarly, 80 source ASes (among several thousand observed) can contribute 95% of the traffic on a link.

In this paper we study the mismatch between number and size in terms of *diversity* and *disparity*. *Diversity* is the presence of a large number of distinct

[3] Support for this work is provided by DARPA NMS (N66001-01-1-8909), DOE Contract No: DE-FC02-01ER 25466, and by NSF ANI-0221172, with support from the DHS/NCS.

objects (e.g., many users sharing a link). Many objects have a natural size measure such as bytes per transfer, customers per provider, and visits per website. *Disparity* is concentration of a size measure in a small subset of objects. For example, a bursty flow may accumulate most of its duration in lulls. In extreme cases of disparity, a *giant cluster* forms, in which the aggregate size is comparable with total volume. We see this with TCP in the IP protocol space, with a popular operating system, and recently, with P2P applications in some networks.

Mathematically, diversity/disparity is present when the counting measure (such as number of addresses) and the size measure (e.g., traffic per address) are close to mutually disjoint (singular), i.e., supported by non-intersecting sets. Many Internet measures of interest are disjoint (e.g., in the case of lulls and bursts, most bytes are transferred in negligible total time). The ubiquity of disjoint measures renders comparison of IP objects challenging.

Neither diversity nor disparity are good or bad per se; the impact depends on the situation. Motivations to study disparity include offsetting its negative impacts (such as lack of resilience) and developing ways to manage total volume via control of a few contributors [1]. Data reduction is another motivation, e.g. frequent objects are assigned shorter bit strings (as in Huffman encoding). Indeed, a valid motto of Internet science is: *Find Disparity in Diversity.*

Size disparity is often called the "mice-elephants" phenomenon. It was observed early in Internet history that Internet traffic displays favoritism at any given aggregation, i.e., many small contributors and a few large ones[4]. Researchers have described duration 'elephants' (long-lasting flows) [3] and bitrate or burst 'elephants' [4]. We focus on volume elephants, due to ISPs' need for a metric for use in pricing as well as due to bitrate limitations of many links.

Comparing diversity/disparity across multiple datasets, directions, measures, percentile levels and source/destination granularities can easily result in an explosion of numbers. We propose a concise characteristic of disparity that we call *crossover*. We define it as the fraction $1 - c$ of volume accumulated in fraction c of top objects. We justify this metric in Section 3. Object size at the crossover x_c serves as a cutoff between the mice and elephant classes (cf. [5]). We study crossovers both empirically and analytically. We think that crossovers are potentially as useful as while being more descriptive than the 95th percentile currently used in many MIBs and autogenerated reporting software.

The paper is structured as follows. Section 2 discusses motivations, and in Section 3 we compute crossovers for uniform, exponential and Pareto distributions. We assess their validity as models for aggregated traffic by comparing their crossovers with observed values. In particular we find that uniform or exponential distributions have crossovers under 70/30. The Pareto density $Cx^{-\alpha-1}$, $1 \leq x \leq N$ approaches the observed splits of 90/10 for $\alpha = -1$ (Zipf distribution) and $N = 10^{10}$, whereas 95/5 splits with realistic values of N can only be obtained for α in $0 < \alpha < 1$ range.

[4] "1% of those networks were responsible for 70% of the traffic [on NSFNET] for the month of December 1992" [2].

Section 4 presents our empirical analysis of diversity and disparity of IP addresses, prefixes, policy atoms and ASes (measured by bytes and packets) for the longest traces in our Backbone Traffic Data Kit. We find that for most combinations of IP categories (from addresses to ASes), measures (bytes or packets) and datasets, crossover ratios are above 90/10, with many above 95/5. We discuss these results in Section 5 and outline future work and conclusions in Section 6.

2 Preliminaries

Motivation. The advent of Dag monitors [6] [7] has facilitated the capture of packet headers for over a terabyte of IP traffic (Table 1). However, humans cannot use such a volume of data in raw form. One needs to reduce 12 orders of magnitude (10^{12} bytes) to one order (a ten-line report) before it can be handled by a person[8]. Fortunately, the size disparity in network data makes reporting "heavy-hitters" possible and useful. We introduce the notion of *crossover* in order to concisely quantify disparity of large data sets. Crossover also allows for a a rough estimate for the number of objects taken into account by a routing/traffic engineering optimizer.

The economic motivation for studying disparity of aggregated traffic can be explicit (e.g., for price differentiation [9]) or implicit, e.g., for security [10], QoS or traffic engineering [1] [5].

Routing-based aggregation. Lots of meaningful aggregations exist between an individual bit and all traffic observed in an experiment. We use 5 categories based on routing (including ports that route data through end systems): flow (source/destination address and port, protocol, 64s-timeout [11]), IP addresses, prefixes, policy atoms [12] and Autonomous Systems (ASes).[5] Routing-based aggregation of IP traffic makes sense because it follows ISPs' income flow.

Measures. The most commonly used measure is the counting one; it assigns 1 to each object. Data aggregation maps one category to another, e.g., bytes to packets, or prefixes to origin ASes. The counting measure is then collapsed to a "marginal" that counts the number of elements in an aggregated object, e.g., bytes per packet or addresses per prefix. It is notable that many measures used in practice map to node degree in graphs. An n-level aggregation hierarchy will produce $n(n-1)/2$ marginal measures, unmanageable even for small n. The multitude of available measures can explain why researchers sometimes derive inconsistent answers to the same question. Disjoint measures (Section 1) can calibrate the concepts of rare, typical and prevalent in divergent, incompatible ways. To avoid the perils of ambiguity, we only use two measures, bytes and packets, for each type of routed object.

Crossover. Knowing the boundary between mice and elephants is a requirement of many traffic engineering schemes [5] [14]. However, there is no natural boundary in the size spectrum. In fact, there are cases where most volume comes from midrange ("mules" [15]) contributions.

[5] We separate source and destination rather than working with a matrix (cf.[13]).

To address these concerns we studied the dependence of the cutoff on the proportion of traffic in the elephant class and aggregation level, and compared the object count median to the size median. We found however that the mice-elephant cutoff is best placed at the crossover point. We are now ready to discuss the virtues of this statistic in detail.

3 Theory.

Mice-elephant boundary. Labeling an object as mouse (contributing to numbers) or elephant (contributing to mass) can be cast as hypothesis testing [16]. Let $f(x)$ be the density of objects of size x (the number of objects of size x divided by the number of observed objects). The mass density at x is $xf(x)/\bar{x}$ where $\bar{x} = \int_0^\infty xf(x)dx$ is the average object size. The null hypothesis H_0, "the object is a mouse," has likelihood function $f(x)$, whereas $xf(x)/\bar{x}$ is the likelihood for competing hypothesis H_1, "the object is an elephant." The maximum likelihood decision would amount to a cutoff at $\bar{x}$: elephants are objects whose size is above the average. This is the point where two densities intersect, ($f(x) = xf(x)/\bar{x}$ at $x = \bar{x}$.) Figure 1(a) shows an example for flow and byte measures' densities. The 10 kb intersection agrees with the average in Table 2 (Section 4). We leave this approach for future work.

Another option is to equalize the type I error (fraction of objects over the cutoff) with the type II error (fraction of mass under the cutoff). This is a natural choice when the cost of each error is unknown, and it follows a statistical tradition of taking confidence intervals with equal significance at either side. We call s the crossover threshold if the share of objects above s and share of mass below s equals c. The proportion $1 - c : c$ is the *crossover split*. Figure 1(b) presents an example of a cdf and ccdf intersection for prefix volumes in D04 (see Section 4.)

An equivalent definition for the crossover s is the point where the cdf of objects crosses the volume ccdf, or where the sum of two cdfs or two ccdfs equals 1. Since the cdf sum monotonically increases from 0 to 2, the crossover always exists. We take the size for which the volume share of top objects and their counts' share add up closest to 100%, which locates the crossover between two medians, for object count and for volume, which serve as upper and lower bounds on the elephant cutoff (Fig.1 (b)).

We can now compute examples of crossovers for some standard distributions. Note that linear transforms $x \mapsto kx$ do not change the crossover split (although shifts $x \mapsto x + a$ do), so we can take any scale factor in the formulas.

For a continuous density $f(x)$, crossover size s satisfies an equation

$$\int_0^s f(x)dx + \int_0^s xf(x)dx/\bar{x} = 1. \qquad (*)$$

We give explicit values for several special cases below.

All sizes equal: object counts equal object sizes, resulting in a 50/50 split.

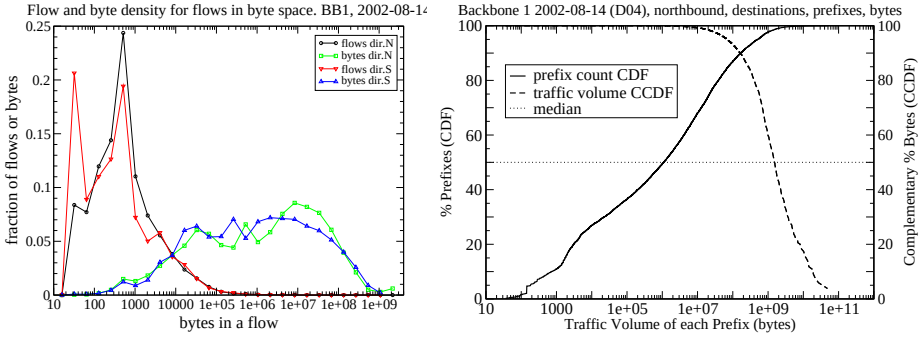

Fig. 1. a) Disjointness of flow and byte measures on 5-tuple flows, D04 b) Crossover for destination prefix byte counts, D04N

Uniform distribution. For $0 \leq s \leq 1$, the equation $s + s^2 = 1$ is satisfied by the *golden section*[16][6] $\frac{\sqrt{5}-1}{2}$. The split is 61.8/38.2, about 60/40.

Exponential: $P(X < s) = 1 - e^{-s}$. Equation (*) reduces to $1 - e^{-s} - se^{-s} + 1 - e^{-s} = 1$, i.e. $2 + s = e^s$; $s = 1.146$ and the split is 68.2/31.8, about 70/30.

Pareto distribution is usually truncated at some N, which specifies the maximum object size. On the interval $[1, N]$, the cdf $F(s) = P(X \leq s) = \frac{1-s^{-\alpha}}{1-N^{-\alpha}}, \alpha > 0$; mass fraction is $M(s) = \frac{1-s^{-\alpha+1}}{1-N^{-\alpha+1}}, \alpha \neq 1$ and Equation (*) becomes

$$\frac{1 - s^{-\alpha}}{1 - N^{-\alpha}} + \frac{1 - s^{-\alpha+1}}{1 - N^{-\alpha+1}} = 1. \qquad (**)$$

For $\alpha = 1$ (Zipf distribution) $M(s) = \ln s / \ln N$.

We derive an approximate solution to $F(s) + M(s) = 1$, or $F(s) = 1 - f$, $M(s) = f$, by dropping $N^{-\alpha}$ from $F's$ denominator. Zipf's 80/20 split is at $N \approx 3000$ $(e^8), s \approx 5$ $(e^{1.6})$ and 90/10 at $N = 10^{10}, s = 10$. The 80/20 rule may thus have its origin in the Zipf distribution with a ratio of 3000 from highest to lowest value. It may well resemble the bitrate consumption disparity found in most ISPs; only a few of them offer a wider range of bitrates (for example, 3.5 orders covers DSL at 512 Kbps to OC-48 at 2.5 Gbps).

When $\alpha > 1$, Equation (**) reduces to $s^\alpha = s + 1$ for large N; for $\alpha = 2$ this again produces a golden section split of 62/38. In general, as α grows, the split evens out since large sizes become less likely.

More extreme splits than 90/10 only occur for Zipf when N exceeds the current typical traffic range (e.g., 95/5 requires $N = 10^{26}$), but Pareto with α outside the conventional interval $\alpha \geq 1$ ([18]) can produce these splits for moderate sizes of N. For $\alpha = 0.5$, for example, Equation (**) holds whenever $s = \sqrt{N}$, resulting in a 90/10 split at $N \approx 6600(e^{8.8}), s \approx 90$, and 95/5 at $N \approx 133,000(e^{11.8}), s \approx 365$. Section 4 confirms these results.

[6] Also known as *golden ratio* or divine proportion [17].

Table 1. Bulk sizes of OC-48 and OC-12 datasets

Set	Bb	Date	Day	Start	Dur	Dir	Src.IP	Dst.IP	Flows	Packets	Bytes
D04N	1	2002-08-14	Wed	09:00	8 h	Nbd (0)	2124 K	4074 K	106.6 M	2144 M	1269 G
D04S	1	2002-08-14	Wed	09:00	8 h	Sbd (1)	1122 K	12661 K	193.8 M	3308 M	2140 G
D05I	U	2002-08-14	Wed	08:22	13 h	Inbd (1)	961 K	11183 K	37.6 M	538 M	326 G
D05O	U	2002-08-14	Wed	08:20	16 h	Obd (0)	25.6 K	1412 K	22.0 M	549 M	249 G
D08N	1	2003-05-07	Wed	00:00	48 h	Nbd (0)	3902 K	8035 K	275.5 M	4241 M	2295 G
D09N	2	2003-05-07	Wed	10:00	2 h	Nbd (1)	904 K	2992 K	56.7 M	930.4 M	603 G
D09S	2	2003-05-07	Wed	10:00	2 h	Sbd (0)	466 K	2527 K	47.3 M	624.2 M	340 G

Table 2. Rates and size ratios for OC-48/OC-12 datasets. New sources, destinations and flows per second are bulk averages over the whole trace (in units of 1000/sec).

Trace	Sr/s	Ds/s	Fl/s	kpps	Mbps	Ut.%	Fl/Sr	Fl/Ds	Pk/Fl	Pk/Ds	Bt/Fl	Bt/Pk
D04N	74	141	3700	74	352	14.2	50	26	20	526	11906	592
D04S	39	440	6730	115	594	23.9	173	15	17	261	11041	647
D05I	21	245	821	12	57	9.2	39	3	14	48	8668	605
D05O	0.44	25	381	10	35	5.6	857	16	25	389	11347	454
D08N	23	47	1594	25	106	4.3	71	34	15	528	8331	541
D09N	126	415	7881	129	671	26.9	63	19	16	311	10635	649
D09S	65	351	6566	87	378	15.2	101	19	13	247	7193	545

4 Diversity and disparity in high-speed traffic

Data sources. We use four longest datasets from our Backbone Traffic Data Kit (Tab. 1): D04, D05, D08 and D09.

The data in D04, D08, D09 was collected by OC-48 monitors using DAG 4 cards from U.Waikato/Endace [7] on Linux platform. D04 contains 8 hours of OC-48 traffic at up to 28.5% utilization (over 1 sec intervals)) taken at Tier 1 Backbone 1 (BB1) in August 2002. D08 and D09 were captured in May 2003. D08 covers 48 hours of Backbone 1 traffic from the same link as D04, albeit at lower utilization. (Only the northbound direction was captured.) D09 contains 2 hours (overlapping with D08) at up to 30.6% OC-48 utilization taken in Tier 1 Backbone 2 (BB2). Backbone links connect San Jose in the south to Seattle in the north. BB1 and BB2 use Packet over Sonet (POS). BB2 also prepends 80% packets with 4-byte MPLS headers (fewer than 50 distinct labels used on each direction; always one label in a stack.) Both encapsulations result in a small reduction of available Sonet payload.[7]

D05 was collected on an OC-12 (622 Mbps) ATM link that carries traffic between a university and the Internet. We label link directions by prevalence of inbound (I) and outbound (O) traffic, although due to multihoming each direction carries both. D05 is taken on the same day as D04. All traces are

[7] The reduction depends on average packet size and the extent of HDLC byte stuffing [19]. Knowing these factors is essential for precise bitrate estimates. Utilizations here and in Table 2 give IP packet volume divided by Sonet raw bitrate of 2488.32 Mbps.

Table 3. Geographic Distribution of Traffic. D04, BB1, August 2002

	north,src	north,dst	south,src	south,dst
amer	77.5% (9.84e+11)	88.7% (1.13e+12)	91.4% (1.96e+12)	64.3% (1.38e+12)
asia	22.4% (2.84e+11)	10.4% (1.32e+11)	8.6% (1.84e+11)	29.4% (6.30e+11)
euro	0.0% (3.94e+08)	0.3% (3.84e+09)	0.1% (1.20e+09)	4.4% (9.44e+10)
other	0.0% (4.87e+08)	0.6% (7.84e+09)	0.0% (8.45e+07)	1.9% (4.05e+10)

Table 4. Geographic Distribution of Traffic. D09, BB2, May 2003

	south,src	south,dst	north,src	north,dst
amer	98.2% (3.34e+11)	77.0% (2.62e+11)	68.8% (4.15e+11)	96.9% (5.85e+11)
asia	1.4% (4.90e+09)	5.0% (1.69e+10)	30.4% (1.83e+11)	3.1% (1.86e+10)
euro	0.0% (3.71e+06)	14.7% (5.01e+10)	0.5% (2.75e+09)	0.0% (6.61e+06)
other	0.4% (1.26e+09)	3.3% (1.13e+10)	0.3% (2.03e+09)	0.0% (2.10e+07)

captured in the middle of the week around noon. We used CoralReef [13] and other CAIDA programs for data processing.

Our backbone data has high geographic diversity (all continents are represented, including Latin America and Africa). However, volumes are dominated by a giant cluster of North American traffic, which comprises at least 2/3 of all traffic in those datasets with significant traffic: D04 (Table 3), D05, and D09 (Table 4). The underutilized dataset D08 consists of half Asian and half American traffic. Asian traffic figures prominently in our other datasets too, likely since the traces were from US West coast links. In the backbone data trace both Asian and European traffic flows south, toward Internet exchanges located in Bay Area.

The raw diversity of our data is high. Traces differ by utilization, traffic symmetry, and temporal dependencies. In Figure 2 each panel shows the number of bytes, flows, source-destination pairs and packets. The remarkable stationarity of the traces with respect to baselines (major bursts are rare) means that volume distributions of these traces are convincing from a stochastic viewpoint.

An interesting property of traces D04 and D09 is equality between orders of magnitude for bitrate (Mbps) and the number of flows active per (64-sec) interval (a flow over 64 sec translates to 1 kbps of average bitrate). The maximum value of flows/second is 48702 for D04S and 20778 for (similarly utilized) D09N (the discrepancy is due to attacks in D04, see below). The maximum number of source-destination pairs is 24310 for D04S (no major attack at this second) and 15531 for D09N. For 64s intervals the maximum numbers of flows is 1.71M for D04S (peak of 220K IP pairs) and 719K flows (peak of 435K IP pairs) for D09N. In addition, D08 and D05 have diurnal variation with a factor of 2-6X.

Another interesting property of our datasets is the almost constant average bytes/flow. Table 2 consistently shows it at around 10 kbytes. Packets/flow, packets/destination, and bytes/packet are also of the same order of magnitude for most traces. The only exception is the inbound university trace D05I, skewed by backscatter [20], scans and other traffic debris attracted to a large address

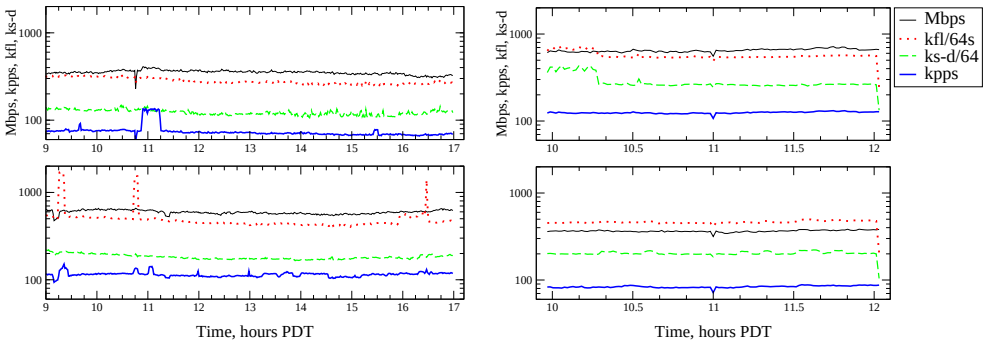

Fig. 2. Traffic rates (64 sec intervals) D04N (left), D09N (right). Byte and packet counts are expressed in Mbps, kpps respectively as per-second averages.

Table 5. AS and Prefix Coverage

Set	Src.Pfx		Dst.Pfx		Src.AS		Dst.AS	
D04N	25.15%	28,208	9.93%	11,141	38.34%	5,288	13.49%	1,861
D04S	7.50%	8,413	39.97%	44,835	14.80%	2,042	45.42%	6,264
D05O	0.43%	487	42.67%	47,858	1.87%	258	72.99%	10,067
D05I	40.96%	45,946	0.54%	615	70.02%	9,657	2.34%	324
D08N	29.51%	33,940	6.94%	7,984	51.98%	7,986	10.57%	1,625
D09N	16.17%	18,595	1.31%	1,510	23.61%	3,628	1.52%	234
D09S	1.52%	3,370	16.78%	19,296	4.07%	626	25.89%	3,978

block.[8] We plan to report in future whether and to what extent these ratios are invariant in high-speed IP traffic.

Prefix/AS diversity. The diversity of prefixes and ASes in our data is also high (Tab.5.) Taken together, sources and destinations on both directions of each link cover 30-55% of RouteViews (semiglobal [22]) prefixes and 42-62% ASes. However, disparity of coverage between directions can be high, e.g., in D09 the northern side of the link has only data from/to 1.5-4% of prefixes and ASes. Another interesting property is the symmetry in coverage disparity. The number of sources on one direction of a link is of the same order of magnitude as the number of destinations on the other, even though these sources and destinations do not necessarily match each other.

Extreme disparity. Packet floods, DDoS attacks and IP address and port scan are usually viewed as traffic anomalies. However, each trace in our study (indeed, almost any wide-area Internet trace) contains examples of all these phenomena. This observation renders 'anomalies' normal in highly multiplexed traffic (cf. similar observations for backscatter data [20]). We view them as cases of extreme size disparity at particular aggregation levels.

In particular, *floods* are typically aimed at overwhelming the capabilities of the receiving machine at the other end, e.g. OS interrupt processing. To reach that goal, lots of small packets are sent. As a result, *packet rate* increases without proportional growth in utilization; cf. curve in the upper left plate of Figure 2.

[8] Our full analysis [21] skips all traffic directed to that block.

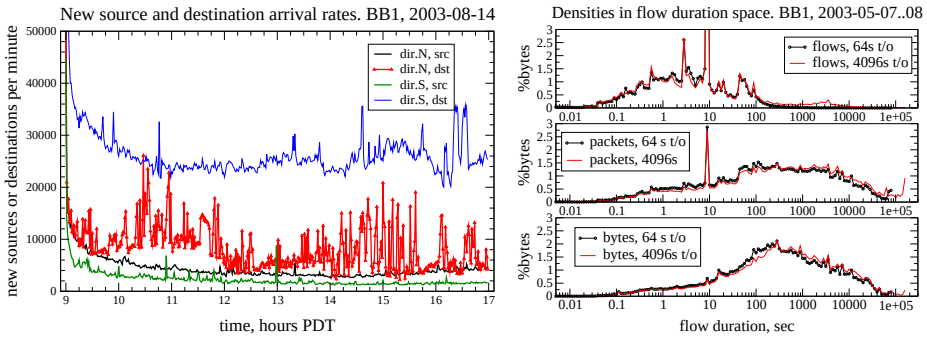

Fig. 3. a) Bursts in destination arrival rates, D04 b) Flows, packet and byte densities in duration space for two timeouts

DDoS attacks represent a particular type of flood; two such attacks appear present on southbound D04 (Figure 2 lower left). The source addresses sweep the whole /16 address blocks of an academic network in Asia; the destination addresses point to hosts on consumer networks in US and Turkey. Note that the excursions at the flow aggregation level are missing on the level of source-destination pairs (because of the restriction to /16s), while more aggressive address spoofing would transpire to that level as well. These attacks also change packet rates (bottom curve), but to a much smaller extent.

Another type of disparity derives from IP address and port *scans* done (among others) by viruses and hackers. This activity appears as a large number of source-destination pairs (address scans) or flows (port and address scans). Figure 3 shows multiple bursts in new source-destination pairs per minute due to repetitive scans going north. Scans may be present in the D09, Fig.2 (right).

Figure 5 (upper left) indicates a large number of destination IP addresses with small (40-200 bytes) traffic volume, concentrated at a few small packet sizes, 40, 40+40, 3*48 bytes, many of which reflect SYN probes from scanning tools. Half of dir.0's destinations and 2/3 of dir.1's destinations favor these packet size modes. Neither source IP cdf's (Fig. 5 left, dotted lines) nor AS distributions (upper right) have this property. Scans also impact packet counts per IP addresses, resulting in large (20-40% for D04 and D09) fractions of destination IPs with only one packet; the corresponding source counts are at most 20%. The impact of scans on traffic volumes per prefix (Figure 5, middle panels) and AS (Figure 5, right panels) is negligible since they affect only one prefix at a time.

The measures of Fig.3(b) (top: flow density; middle: packets; bottom: bytes) reveal the presence of scans in D08 differently. While demonstrating disparity in flow duration (by 7 orders of magnitudes) and the prevalence of *mules* [15], they also have several strong spikes. Many scans send 2 or 3 SYN packets to the same address, retransmitting after standard timeouts of 3 and 6 seconds. As a consequence, many flows (almost 12.4%) have duration around 9 sec (8.9-10.0) and about 2.6% at 3 sec (2.81-3.16). Packet density has a spike at 9 sec, while for byte density it is a barely visible bump.

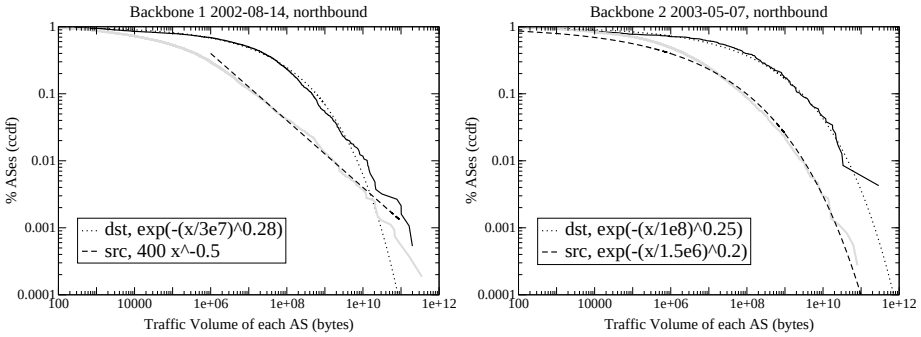

Fig. 4. Ccdf for source and destination bytes per AS. D04N (left), D09N (right).

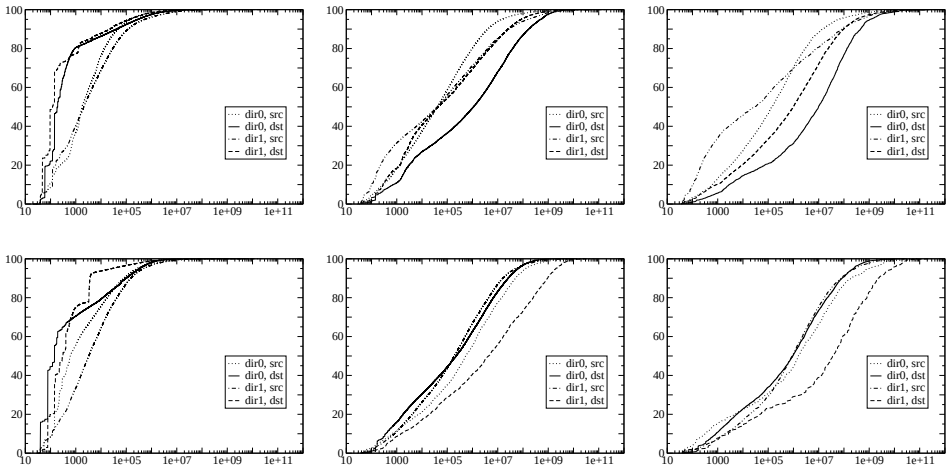

Fig. 5. IP addresses (left), prefixes (center) and ASes (right) as traffic sources (dotted lines) and sinks (solid and dashed lines). Object count cdfs over traffic volume (bytes per object). D04 (first row) and D09 (second).

Long tails. As shown in Section 3, the dominance of the contribution of the largest sizes is a consequence of the heavy tail. Internet loads cover many orders of magnitude, and the number of objects decays slowly.

We find that traffic distribution per AS in traces D04 and D09 is closer to Weibull (0.2-0.3) than to Pareto distribution (Figure 4), and even when its tail is close to Pareto, α can be as small as 0.5 (Fig.4, left panel, source ASes)) which results in stronger bias towards elephants a Zipf distribution with the same range would predict. In future work we plan to compare this result to the Pareto approximation with $\alpha \sim 1$ for 5-minute prefix byte volumes in [5].

Diversity at fixed percentiles. 95% of bytes in our data are sourced by fewer than 8% and destined to fewer than 20% of IPs, prefixes, and ASes. We find that the fraction of IP addresses comprising any traffic percentile is always smaller

Table 6. Crossover for D04 (left) and D09 (right). Bytes (top), packets (bottom).

	N,src	N,dst	S,src	S,dst
IP	97.5/2.5	96.2/3.8	97.4/2.6	96.4/3.6
Pf	97.2/2.8	89.9/10	96.6/3.4	93.7/6.3
At	97.1/2.9	92.1/7.8	97.0/3.0	94.5/5.5
AS	97.4/2.6	92.3/7.7	97.1/2.9	95.1/4.9
	N,src	**N,dst**	**S,src**	**S,dst**
IP	93.8/6.2	95.7/4.3	94.7/5.3	95.8/4.2
Pf	94.7/5.3	90.2/9.8	94.9/5.1	93.7/6.3
At	95.3/4.7	92.0/8.0	95.6/4.4	94.3/5.7
AS	96.0/4.0	92.2/7.8	95.8/4.2	94.9/5.1

	S,src	S,dst	N,src	N,dst
IP	97.0/3.0	93.1/6.9	95.6/4.4	97.9/2.1
Pf	93.5/6.4	89.5/10.5	93.6/6.4	87.6/12.5
At	93.4/6.6	89.0/11.0	93.2/6.8	89.4/10.5
AS	94.1/5.7	91.1/8.9	93.9/6.1	90.8/8.9
	S,src	**S,dst**	**N,src**	**N,dst**
IP	94.4/5.6	92.6/7.4	91.1/8.9	97.2/2.8
Pf	90.8/9.2	89.7/10.3	90.1/9.9	87.0/13.0
At	91.3/8.7	89.2/10.8	90.6/9.4	88.9/11.3
AS	92.8/7.2	91.3/8.7	91.9/8.0	90.6/9.4

than the fraction of networks (prefixes, atoms, ASes); the fractions of networks are usually close to each other. More details of the analysis are at [21].

Crossovers. Table 6 quantifies the diversity/disparity of our data in terms of crossovers. We find that crossover splits are far more extreme than 80/20; in fact, almost all of them are in 90/10 range, and many exceed 95/5. This is in particular true for IP addresses, for which the disparity is highest. Packets disparity is usually higher than bytes' for IP addresses, but it is comparable for prefixes, atoms and ASes. Another property that also holds true for fixed percentiles is the position of atoms close to (often between) prefixes and ASes. The analysis in [21] also shows small numbers of atoms and ASes responsible for 50% of all traffic. In particular, the number of ASes responsible for half the traffic is always under 20 for backbone traces D04, D08, D09. These properties make atoms and ASes good candidates for use in traffic engineering [14].

5 Discussion

The phenomenon of size diversity/disparity has multiple causes. One cause strongly suggested by our data is the variable amount of aggregation present in the objects of the same taxonomic level. For example, traffic from an IP address may be generated by a single person, but it can also come from a network behind a NAT, potentially with thousands of addresses. Traffic toward an IP address may be destined to one user or to a popular news server; in the latter case the IP flow incorporates millions of individual contributions transferred as a single network news feed. A large AS with a large set of connected customer ASes may have inherited many of them via mergers and acquisitions. The same applies to other measures of wealth. The nature of the process shaping size distributions is a matter of debate dating back to Yule's 1924 paper [23] (cf. [24].) Indeed, no single model is likely to fit all cases of size disparity, even if limited to long-tailed size distributions. However, our prior work [25] presents evidence that some long-tailed distributions can arise from multiplicative coalescence, a process in which the probability of joining two objects is proportional to a power of their sizes' product.

6 Concluding remarks

The rich hierarchy of categories used in IP traffic analysis yields many aggregated measures that can serve as foundations for differentiating typical from rare and extreme. Many of these measures are mutually exclusive, which can affect research conclusions unless the disjointness, in particular diversity/disparity and similar phenomena, are explicitly considered.

In this paper we suggested size disparity as a unifying paradigm shared by seemingly unrelated phenomena: burstiness, scans, floods, flow lifetimes and volume elephants. We pointed out that in general an aggregated measure has a meaning of node degree in some graph. We then discussed concentration properties of byte and packet measures aggregated by IP address, prefix, policy atom and AS. We found that an attempt to faithfully quantify diversity/disparity in Tier 1 backbone data leads to combinatorial explosion of the parametric space. To reduce the description complexity, we introduced a mice-elephant boundary called *crossover*. We showed that many IP traffic aggregation categories have crossovers above the proverbial 80/20 split, mostly around 95/5. We also found that the Pareto models, previously used for file/connection/transfer sizes [18] and short-term prefix traffic volumes [5], require a significant bent ($\alpha \sim 0.5$) to account for the size disparity of aggregated and accumulated backbone traffic. On the other hand, a Weibull distribution with shape parameter 0.2-0.3 can serve as an alternative model for the tails of AS volume data.

Due to space limitations we could not include all analyses done for this study. Our results, including geotraffic volumes, diversity of objects that contribute over 1% of traffic, consumers of fixed (50, 90, 95, 99) traffic percentile volumes, crossover fractions and cutoffs, volume of mice and distribution plots (all for bytes and packets) are available at [21].

We intend to extend this study of 'volume elephants' to the 'elephants of stability', i.e. flows with low variation that are important for traffic engineering [14] [5]. We plan to investigate disparity measures based on Shannon's entropy notion, and to find which natural category of traffic aggregates can fill the (two orders of magnitude by packet or byte volume) gap between IP addresses and prefixes. Finally we hope that crossovers, introduced here for the first time, will serve as a bridge between academic and operational parts of networking community, combining a mathematically precise value on one side with a familiar concept on the other.

Acknowledgements. Thanks to Joerg Micheel, Nevil Brownlee and Dan Andersen for the Backbone Trace Data Kit, to Ken Keys for analysis software, to Patrick Verkaik for MPLS data, and to Khushboo Shah for dicussions of scans.

References

1. A.Shaikh, J.Rexford, Shin, K.G.: Load-sensitive routing of long-lived IP flows. In: SIGCOMM. (1999)
2. kc claffy: Internet traffic characterization (1994) Ph.D. thesis, UCSD.

3. N.Brownlee, kc claffy: Understanding Internet traffic streams: Dragonflies and Tortoises. In: IEEE Communications. (2002)
4. K.Lan, J.Heidemann: On the correlation of Internet flow characteristics (2003) Report ISI-TR-574, www.isi.edu/trpublic/pubs/au-kclan.html.
5. Papagiannaki, K., Taft, N., Diot, C.: Impact of flow dynamics of traffic engineering principles. In: INFOCOM. (2004)
6. I.Graham, M.Pearson, J.Martens, S.Donnelly: Dag - a cell capture board for ATM measurement systems (1997) wand.cs.waikato.ac.nz.
7. Endace: Measurement Systems (2004) www.endace.com.
8. C.Estan, S.Savage, G.Varghese: Automatically inferring patterns of resource consumption in network traffic. In: SIGCOMM. (2003)
9. Odlyzko, A.M.: Privacy, economics, and price discrimination on the Internet. In: ICEC ACM. (2003)
10. R.Mahajan, S.M.Bellovin, Floyd, S., J.Ioannidis, V.Paxson, S.Shenker: Controlling High Bandwidth Aggregates in the Network. In: CCR. (2002)
11. Claffy, K., Braun, H.W., Polyzos, G.: A Parametrizable methodology for Internet traffic flow profiling. In: IEEE JSAC. (1995)
12. Broido, A., k claffy: Analysis of Route Views BGP data: policy atoms (2001) Proceedings of NRDM, Santa Barbara.
13. Moore, D., Keys, K., Koga, R., Lagache, E., kc claffy: CoralReef software suite as a tool for system and network administrators. In: Usenix LISA. (2001)
14. N.Feamster, J.Borkenhagen, J.Rexford: Guidelines for interdomain traffic engineering. In: CCR. (2003)
15. Broido, A., E.Nemeth, kc claffy: Spectroscopy of private DNS update sources. In: WIAPP03. (2003)
16. Korn, G.A., Korn, T.M.: Mathematical handbook for scientists (1968)
17. R.Knott: Fibonacci numbers and the golden section (2003) www.mcs.surrey.ac.uk/Personal/R.Knott/Fibonacci/fib.html.
18. Crovella, M.E., Bestavros, A.: Self-similarity in World Wide Web traffic. Evidence and possible causes. In: IEEE/ACM Transactions on Networking. (1997)
19. W.Simpson: PPP in HDLC-like Framing. In: RFC1662. (1994)
20. D.Moore, Voelker, G., S.Savage: Inferring Internet Denial-of-Service Activity. In: USENIX Security Symposium. (2001)
21. Supplement: (2004) www.caida.org/analysis/workload/diversity.
22. Broido, A., E.Nemeth, kc claffy: Internet expansion, refinement and churn (2002) ETT 13.
23. G.U.Yule: A mathematical theory of evolution. In: Phil. Trans. Roy. Soc. London Ser. B, 213:21–87. (1924)
24. Mitzenmacher, M.: A brief history of generative models for power law and lognormal distributions. In: Internet Mathematics. (2003)
25. A.Broido, kc claffy: Analysis of routing and topology data (2001) www.caida.org/outreach/presentations/routingtopology2001/.