APPENDIX B: Synthetic Traffic Generation

To generate self-similar traffic, we used the method described in [45], where the resulting traffic is an aggregation of multiple sub-streams, each consisting of alternating Pareto-distributed ON/OFF periods.

Pareto distribution is a heavy-tailed distribution with the probability-density function

$$f(x) = \frac{\alpha b^{\alpha}}{x^{\alpha+1}}, \quad x \geq b, \tag{B-1}$$

where α is a shape parameter (1< α < 2), and b is a location parameter. Pareto distribution with 1 < α < 2 has a finite mean and an infinite variance.

In our implementation, each sub-stream generates packets of constant size, though this size is different for different streams. To achieve the required packet-size distribution (like tri-modal distributions reported in [12] or [30]), some sub-streams have higher relative load than the other sub-streams. Multiplexing (serializing) packets from different sub-streams produces self-similar traffic with the desired packet-size distribution.

Each sub-stream generates packets in groups (packet trains or bursts). The number of packets per burst (ON period) follows the Pareto distribution with a minimum of 1 (i.e., the smallest burst consists of only 1 packet) and shape parameter α = 1.4. The choice of α was prompted by measurements on actual Ethernet traffic performed by Leland et al. [15]. They reported the measured Hurst parameter of 0.8 for moderate network load. The relationship between the Hurst parameter and the shape parameter α is $H = (3-\alpha)/2$ (see [47]). Thus, α = 1.4 should result in H = 0.8.

OFF periods (intervals between the packet trains) also follow the Pareto distribution, though with the shape parameter α = 1.2. We used heavier tail for the distribution of the OFF periods because the OFF periods represent a stable state in a network, i.e., a network can be in OFF state (no packet transmission) for an unlimitedly long time, while the durations of the ON periods are ultimately limited by network resources and necessarily finite file sizes. The location parameter *b* for the OFF periods was chosen so as to obtain a desired load ℓ_i from the given sub-stream *i*:

$$\ell_i = \frac{E[ON_i]}{E[ON_i] + E[OFF_i]} \tag{B-2}$$

where $E[ON_i]$ and $E[OFF_i]$ are expected lengths (durations) of ON and OFF periods of source *i*. To generate Pareto-distributed values, we used the following formula

$$X_{PAR} = \frac{b}{U^{1/\alpha}} \tag{B-3}$$

where *U* is a uniform random variable $(0 < U \le 1)$. But, since computers generate discrete values, we have to consider a truncated-tail Pareto distribution. Let us denote by U^{MIN} the smallest non-zero value of *U*. Then, the largest Pareto-distributed value is $X^{MAX} = \frac{b}{\left(U^{MIN}\right)^{1/\alpha}}$.

We can find the p.d.f. $f_T(x)$ of a truncated-tail distribution as follows:

$$\int_b^{X^{MAX}} f_T(x)dx = 1 \quad \Rightarrow$$

$$f_T(x) = \frac{f(x)}{\int_b^{X^{MAX}} f(x)dx} = \frac{\alpha b^\alpha}{x^{\alpha+1}} \times \frac{1}{1 - \left(\frac{b}{X^{MAX}}\right)^\alpha} = \frac{\alpha b^\alpha}{x^{\alpha+1}(1 - U^{MIN})} \tag{B-4}$$

Then, the expected value of a truncated-tail series is:

$$E[X] = \int_{b}^{X^{MAX}} x\, f_T(x)\, dx = \frac{\alpha b^{\alpha}}{1-U^{MIN}} \times \left.\frac{x^{1-\alpha}}{1-\alpha}\right|_{b}^{X^{MAX}} = \frac{\alpha b^{\alpha}}{\alpha - 1} \times \frac{1-\left(U^{MIN}\right)^{\frac{\alpha-1}{\alpha}}}{1-U^{MIN}} \quad \text{(B-5)}$$

Now, we can find the location parameter for OFF periods b_{OFF}. From Eq. (B-2) we get:

$$E[OFF] = E[ON] \times \left(\frac{1}{\ell} - 1\right) \quad \text{(B-6)}$$

After substituting Eq. (B-5) into (B-6), we have

$$\frac{\alpha_{OFF} b_{OFF}}{\alpha_{OFF} - 1} \times \frac{1-\left(U^{MIN}\right)^{\frac{\alpha_{OFF}-1}{\alpha_{OFF}}}}{1-U^{MIN}} = \frac{\alpha_{ON} b_{ON}}{\alpha_{ON} - 1} \times \frac{1-\left(U^{MIN}\right)^{\frac{\alpha_{ON}-1}{\alpha_{ON}}}}{1-U^{MIN}} \times \left(\frac{1}{\ell} - 1\right) \quad \text{(B-7)}$$

From Eq. (B-7), we find b_{OFF}:

$$b_{OFF} = b_{ON} \times \frac{\alpha_{ON}}{\alpha_{OFF}} \times \frac{\alpha_{OFF} - 1}{\alpha_{ON} - 1} \times \frac{1-\left(U^{MIN}\right)^{\frac{\alpha_{ON}-1}{\alpha_{ON}}}}{1-\left(U^{MIN}\right)^{\frac{\alpha_{OFF}-1}{\alpha_{OFF}}}} \times \left(\frac{1}{\ell} - 1\right) \quad \text{(B-8)}$$

With our default values $b_{ON} = 1$, $\alpha_{ON} = 1.4$, $\alpha_{OFF} = 1.2$, and $U^{MIN} = 2^{-32}$, we get:

$$b_{OFF} \approx 0.597 \times \left(\frac{1}{\ell} - 1\right) \quad \text{(B-9)}$$