

From a NetworkWorld article in 2012

<http://www.networkworld.com/article/2186765/lan-wan/arista-10g-switch--fast-and-flexible.html>

Buffering capacity

Arista requested that we measure the burst-handling characteristics of the 7508, specifically to verify Arista's claim that the system can buffer up to 50MB per port. Handling short, high-speed bursts of traffic is especially important in many high-performance computing applications, where multiple senders may present data to the same receiver at the same instant.

While many vendors talk about microbursts in marketing collateral, there isn't yet an industry-standard method of measuring burst handling. We used a couple of methods here: First, with a 2:1 oversubscription of steady-state traffic, where we offer traffic to 256 ports, destined to all the remaining 128 ports. That's a simple buffer test and should work regardless of burst length.

Second, to assess microburst buffering, we sent bursts of varying sizes at line rate from multiple sources to the same destination port at the same time. By experimenting with different burst lengths, we found the maximum microburst length the system could buffer without frame loss.

While the microburst method is arguably more interesting due to the dynamic nature of enterprise traffic, the first method produced a surprising result.

Faced with a 2:1 oversubscription, the switch initially dropped nearly 60% of traffic rather than the expected 50% or less, meaning it wasn't buffering at all. Arista attributed the loss to a combination way the 7508's virtual output queuing (VOQ) works and the totally nonrandom order of our test traffic. After setting the VOQ scheduling to a non-

default setting ("petra voq tail-drop 2"), packet loss fell to 50% or less, as expected.

Another lesson learned, both in steady-state and microburst buffering tests, is that buffer capacity depends in part on the number of senders and receivers involved. When we ran the microburst test with 256 transmitter and 128 receiver ports, the 7508 buffered up to 83.49 megabytes on each receiver port with zero frame loss, well in excess of Arista's claim of 50MB/port. That's equivalent to around 56,300 1,518-byte frames.

However, if we ran the same test with 383 transmitters all aimed at one receiver, the largest amount of traffic that could be buffered without loss was much lower, around 6.85MB (or around 4,600 1,518-byte frames).

The results differ because of the 7508's VOQ and credit-based architecture. When frames enter the switch, it will allocate buffers and issue forwarding credits if, and only if, sufficient resources exist to forward the traffic. The higher the ratio of transmitters to receivers, the greater the imbalance between requested and available resources. In this light, Arista's 50MB claim is really a composite figure, one that assumes transmit and receive port counts are somewhere between the best- and worst-case scenarios.