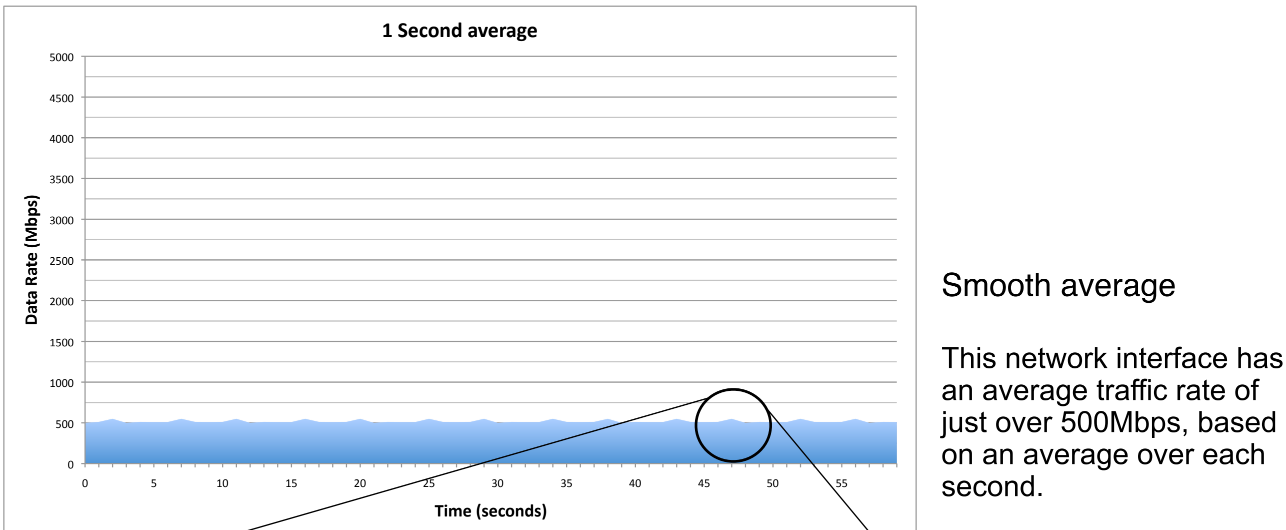
What are Microbursts?

Network traffic is often measured using the average utilization of a link – 5 minute input or output rate, expressed in Mbps or Gbps. The 5 minute averages, or even the 1 second averages are usually smooth – showing the steady state of the network.

The actual traffic in a network when viewed at a finer granularity (such as every millisecond) is far more bursty. These bursts are so fine, that they are often missed by standard monitoring tools. Microbursts are these short spikes in network traffic.

Figure 1: Average traffic on a port based on one second averages



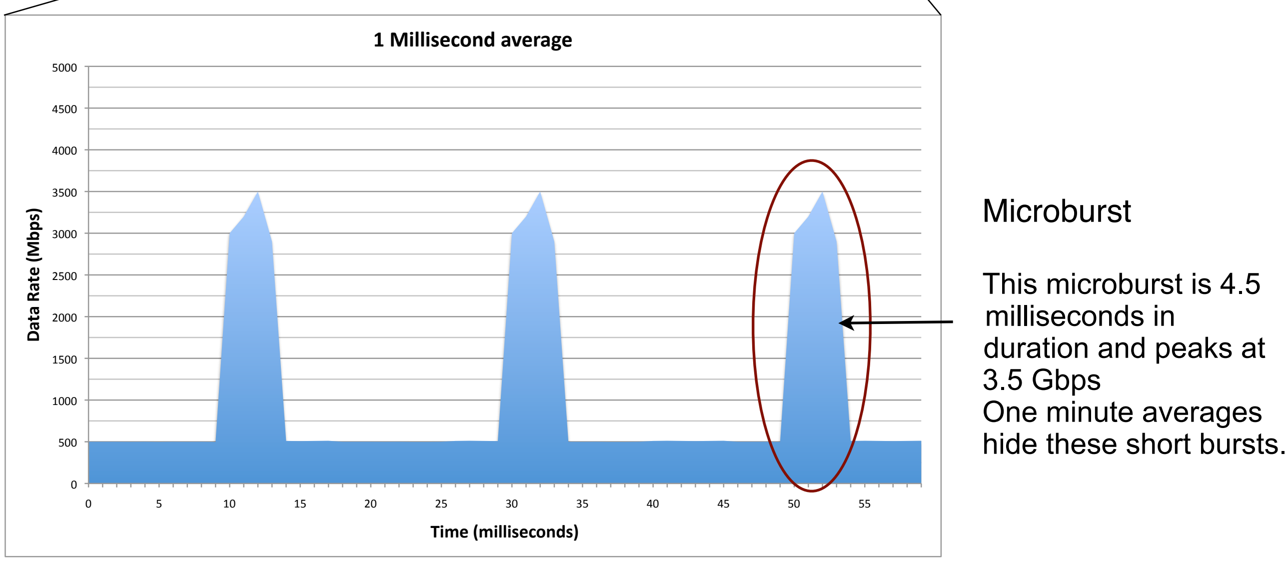


Figure 2: Microbursts – seen with finer time resolution

What causes Microbursts?

Network traffic originates on servers or clients and gets forwarded through multiple hops. Often servers handle multiple client connections and each client gets a slice of the server's compute and I/O cycles. As a result, while the overall throughput of a server may appear to be smooth, traffic within each client session is very bursty.

In addition, a sudden spike in network traffic can result in queueing of packets. That increases jitter in the traffic and that has a ripple effect through other elements of the network.

In the late 1990s, microbursts were a nightmare to deal with in the financial trading applications. Networks were just getting migrated from 100Mbps to 1Gbps. Even with 1Gbps, a burst of data could swamp a traders workstation and impact their trading activity. However, those days are long gone as financial companies have purpose built trading platforms using low latency 10GbE networks.

Jitter due to Microbursts

"Jitter" is the variance in switching latency for packets in a data stream. One very visible impact of microbursts is increased jitter in traffic as it gets forwarded through various network elements. The burstiness of the traffic, coupled with fan-in can result in queuing, which results in increased switching delays. However, even when there is no fan-in, traditional switching asics have slow clock rates and show very high jitter with bursty traffic. While it gets hard to quickly test with specific bursts, lab tests can easily be built to test switches to measure latency with random sized packets instead of fixed size, smooth streams. The quick variation in packets lengths mimics real-world scenarios and the higher jitter is visible as packets are only forwarded when aligned to specific clock cycles.

Packet Buffers to handle Microbursts

Several improvements have been made to switching architectures to forward packets with minimal delay. Cut-through switches do not store the packets in most cases and can forward them in less than a microsecond. However, if congestion builds up, packets need to be queued within a switch. Some switches have a few Kilobytes of ingress packet memory per port. These architectures can handle small amounts of congestion when all ports are congested at the same time. In most applications though, the packet memory is wasted as it cannot be shared by other ports that need to store packet data as queues build up.

Other switches have more packet memory that is external to the forwarding asics. However, external memories require extra clock cycles for lookup and as a result, can only operate in store-and-forward mode and hence have higher latency.

Dynamic Buffer Allocation

So how does one go about getting the lowest latency without compromising the ability to handle congestion? The dynamic buffer allocation scheme used on Arista 7100 series switches provides the solution. Each switching asic provides 2MB of "on-chip" packet buffers that are dynamically allocated to the ports that need it during congestion. As a result, packets get forwarded under a microsecond in the normal case. In case of congestion, each port can use the dynamic memory to store packets, until the congestion clears up. The dynamic allocation of packet memory results in no wastage of memory resources as it only needs to be shared across congested ports.

Summary

Microbursts are short spikes in network traffic that result in higher jitter and sometimes congestion in a network. Legacy switching architectures with slow clock rates exhibit high jitter with such traffic patterns. In addition, packet buffers need to be allocated appropriately to handle congestion in order to avoid drops.

Arista 7100 switches use asics that are clocked at very high rates and can forward packets in less than a microsecond. There is also very low jitter induced due to microbursts. The dynamic buffer allocation in the switches provides a unique approach to managing congestion and having more memory available when needed, compared to other architectures that have per-port, fixed packet memory.