A Primer on Real World High Frequency Trading Taxonomy

Inside

Why
Several vendors have been claiming 'real world' testing in the financial markets based on obtusely contrived 'synthetic tests'

Who Cares
Trading technologists who want the most competitive execution of their trading strategies

What Is Next
Arista will continue to lower the latency and deliver the measurement and management capabilities necessary to compete in the high frequency trading market.

Test multiple vendors, learn what the highest volume firms are doing, and explore how on the days with the highest market volumes Arista has not dropped a frame, or delayed a transaction.

This document was written to inform and educate about the realities and misconceptions of real-world financial trading networks, how to maximize the competitive advantage of your network, and to assist in navigating the world of network device testing methods and terms.

RFC2544 - an informational Request For Comments published in March of 1999. It was authored by Scott Bradner, an Internet luminary from Harvard University and John McQuaid from NetScout. It has been the de facto standard for measuring the performance of network equipment and ensuring a fair baseline of equipment capabilities across vendors for the past decade.

In recent months, RFC2544 has come under fire by some vendors when their products did not win tests based on it. They are advocating for more "real world'" testing and not the same interoperable standard that has sufficed, unmodified, with no other vendor contribution or recommendations for changes since it was authored eleven years ago.

Buffers- very high speed memory used to store packets when a network device is experiencing congestion. The two main causes of congestion are when an ingress interface is transmitting data faster than the egress interface can transmit, and when multiple ingress interfaces transmit concurrently to one egress interface. Buffering and queuing are the leading sources of latency in an Ethernet switching system. However, buffers are required when you experience congestion so that traffic is not discarded forcing a retransmission or causing loss. The sizing and allocation of buffers is a very important factor, requiring balance between congestion management and ultra low-latency.

Microbursts- Microbursts are created when the amount of egress bandwidth on an interface is overrun by the incoming packets sent to that interface for a very short period of time, typically microseconds. Over a longer period of time, this same event would be simply
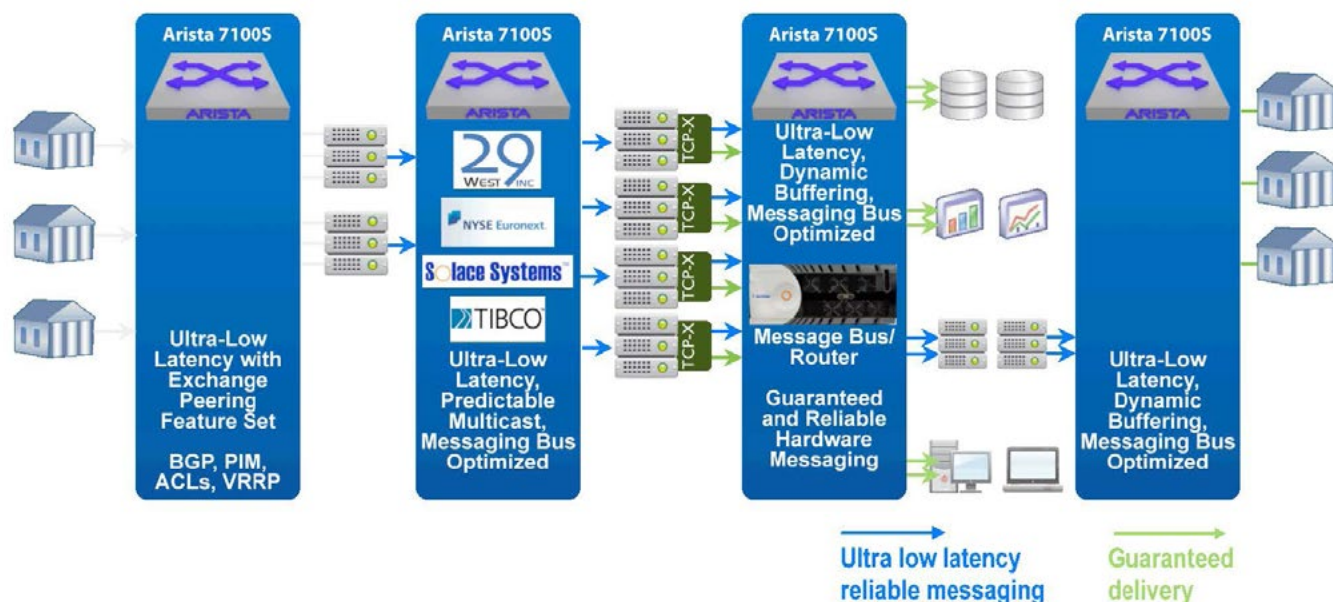
Sample Exchange Peering Interconnect and Electronic Trading Architecture

considered "congestion", or "oversubscription". As a consistent stream of traffic approaches the upper reaches of the overall interface bandwidth, the risk of microburst based congestion or loss increases.

In legacy financial market data installations, "microbursts" caused buffers to fill, packets to be delayed, and potentially dropped. Most of these networks used store and forward Fast Ethernet or Gigabit Ethernet infrastructure. With the advent of wire-speed 10 Gigabit cutthrough Ethernet switching, the available bandwidth now far outstrips the real world market data requirements. This minimizes the risk of event based microbursts effecting latency or packet delivery.

For example: On the BATS financial exchange cross connect, 86Mb/s is the raw feed average bandwidth measured over a one-second average. However, when measured at a onemillisecond time - interval the BATS feed bandwidth can reach 382Mb/s. A 382Mb/s burst might concern those with Gigabit Ethernet cross connects, but a 10 Gigabit link provides exponentially more available bandwidth than is required for even the highest message volumes seen in the real world today.

The difference between Fan-In and Fan-Out - IP Multicast is a commonly used technology that was designed for "few to many" IP data transmission and is used prominently in financial market data feed networks. Some vendors have misrepresented the use of IP multicast to create tests that show some large number of interfaces (23 or 47 are common numbers) talking to one interface via IP multicast, where all the interfaces transmit at maximum rate for a short time interval. This is then referred to as a "microburst". In reality, this pattern of "many to one" multicast traffic is counter-intuitive, and would ONLY be used in a contrived manner, disguised as a multicast test and never seen in the real world.

This confusion of "fan-in" versus "fan-out" multicast is a predictable 'smoke and mirrors' tactic. It is a synthetic test that is carefully created to point out the potential deficiencies of another company's product.

This confusion of "fan-in" versus "fan-out" multicast is a predictable 'smoke and mirrors' tactic. It is a synthetic test that is carefully created to point out the potential deficiencies of another company's product. Usually the aggressor in these cases has a sub-standard product, so they cannot or will not offer up honest standard performance measurements of their products based on the aforementioned RFC2544 metrics. It also obfuscates the real world deployment model of multicast, and the most typical deployment in real world financial networks.

Questions to ask:

1) Some have claimed that tests of 23 interfaces of multicast traffic fanning IN to one interface is a "real world" traffic pattern in high frequency trading. Have you ever seen a 23:1 fan-in in any electronic trading network? How is this test indicative of real-world performance or deployments?

2) If some of the highest volume market data feeds transmit data in a 1 millisecond time-slice at 382Mb with an average throughput of 86Mb/s on a onesecond time-slice, how do you fill any buffer at 10Gb with a 382Mb/s max burst rate?

3) Perhaps your trading strategy is getting an order to the matching engine before another entity at a given price or event. The consistent LOWEST latency throughout the day helps your trading strategy be more competitive. If during a microburst your network switch delays traffic up to 2 milliseconds due to congestion, is it likely the host or feed handler receiving the data will mark it as stale based on time-stamping, an perhaps will drop it? If this was an execution, would you want your message to arrive milliseconds behind?

Santa Clara—Corporate Headquarters
5453 Great America Parkway,
Santa Clara, CA 95054

Phone: +1-408-547-5500
Fax: +1-408-538-8920
Email: info@arista.com

Ireland—International Headquarters
3130 Atlantic Avenue
Westpark Business Campus
Shannon, Co. Clare
Ireland

Vancouver—R&D Office
9200 Glenlyon Pkwy, Unit 300
Burnaby, British Columbia
Canada V5J 5J8

San Francisco—R&D and Sales Office
1390 Market Street, Suite 800
San Francisco, CA 94102

India—R&D Office
Global Tech Park, Tower A & B, 11th Floor
Marathahalli Outer Ring Road
Devarabeesanahalli Village, Varthur Hobli
Bangalore, India 560103

Singapore—APAC Administrative Office
9 Temasek Boulevard
#29-01, Suntec Tower Two
Singapore 038989

Nashua—R&D Office
10 Tara Boulevard
Nashua, NH 03062