The Next Generation Trading Infrastructure

What You Will Learn

Optimizing latency and throughput is essential to high performance trading companies. They are challenged, however, by the complexity of their systems and the need to integrate state-of-the-art technology. This white paper, with validated latency for NASDAQ ITCH market data, shows that this challenge can be overcome with a robust ecosystem of vendors, the right use of technology and an integrated, end-to-end, customized architecture.

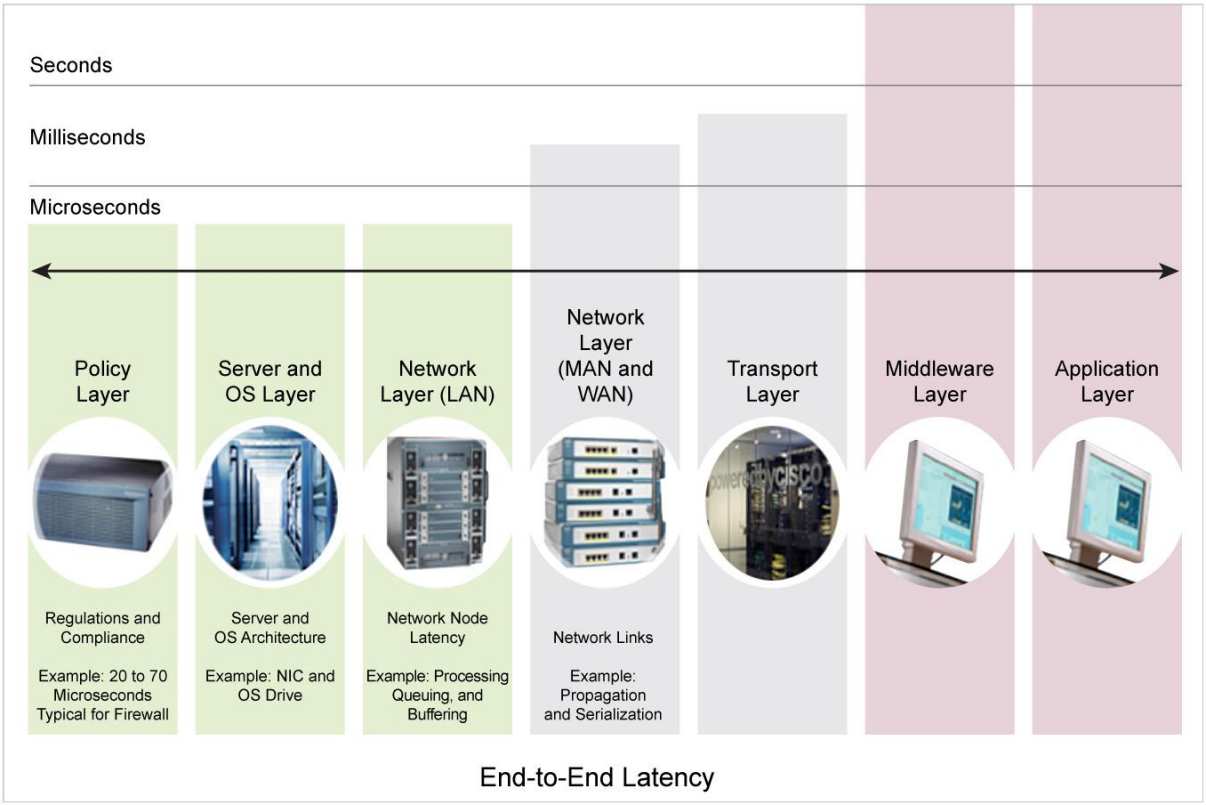
The State of High-Performance Trading Architecture

For HFT and other high performance trading architectures, infrastructures with near-zero latency are essential. The rapid expansion of automated and algorithmic trading has increased the critical role of low latency, high throughput network and server technology needed to process the high volume of transactions. To meet these requirements, low-latency, deterministic system performance is now a pre-requisite characteristic in the financial services market to support HFT, market data distribution, and exchange data processing. In these applications, low-latency networking is critical for processing arbitrage transactions and algorithmic trading with little delay.

In the fast-paced world of high performance trading, lucrative opportunities such as liquidity imbalances and short-term pricing inefficiencies may only exist for microseconds. Success depends on identifying and acting on these opportunities more quickly than the competition. The quest to increase speed has focused on different parts of the trading infrastructure. Field-programmable gate arrays (FPGAs), kernel bypass, switching, TCP offload, and other components have improved overall trading performance. Today, the goal for high performance trading networks is ultra-low latency. Latency capabilities that were impossible a short time ago are now common.

Organizations seeking to reduce latency further now need look no farther than their own co-located equipment. There, simply replacing a traditional router with an ultra-low-latency switch, such as the new Cisco Nexus® 3548 platform switches, can make a significant difference.

Figure 1: Latency Sources



The proof-of-concept (PoC) example described in this document illustrates two main advances in low-latency architecture and their integration into the overall design.

- Network switch latency: Traditionally, network switch design and features such as Network Address Translation (NAT), routing, and policy have resulted in latency between ports or between network nodes in the range of 0.5 to 1.2 microseconds. The introduction of the Cisco Nexus 3548 Switch eliminates this bottleneck with a consistent latency of 50 to 250 nanoseconds (ns) using a unique system-on-a-chip (SoC) innovation.

- Application latency: One of the main advances in high performance trading technology has been FPGA-based processing on the server-side network interface card (NIC) to enable ultra-fast processing of various trading applications such as data acquisition, normalization, risk matching, and order processing. In the market two major design choices is debated - distributed vs. centralized FPGA processing. The centralized FGPA designs have efficiencies in consolidation and single point of aggregation; however distributed FPGA processing has distinct advantages over centralized. Trading architect must consider following criteria while selecting advances in FPGA based design:
  - Technology advances allow faster refresh cycles with distributed design, without upgrade of network and FPGA components. Essentiality allows component based optimization
  - Deeper instrumentation and ties to applications and user space since the infrastructure of application is well developed
  - Scaling and optimized flow from tick to trade. For an example ability to run multiple vendor's IP in multiple servers and scaling for consolidated run book while keep the latency hop to minimum
  - Operational complexity of lifecycle of managing both network switch and FGPA while providing redundancy, software upgrade and operation domain management (network vs. application)

Proof-of-Concept Overview

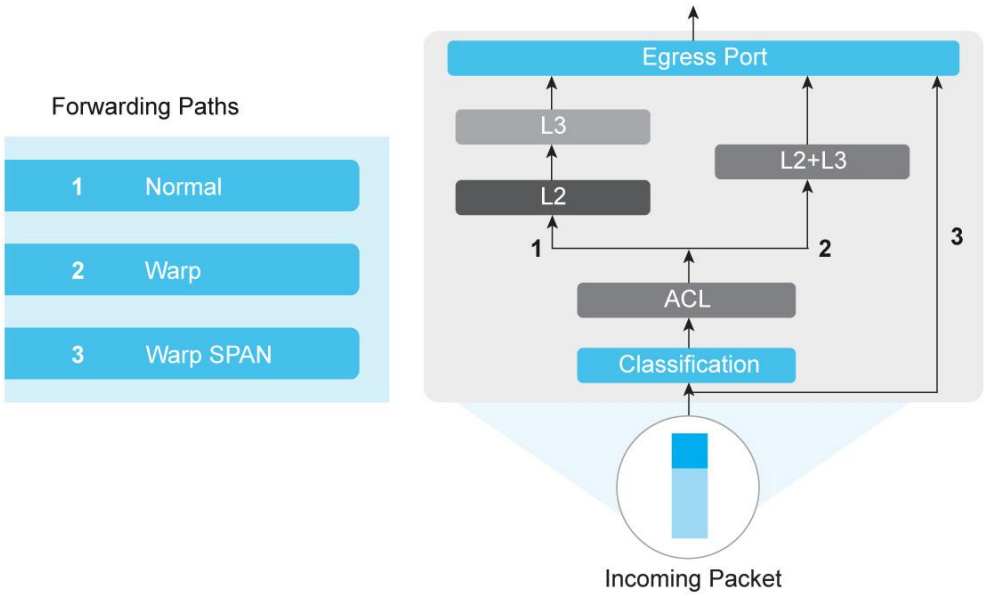
This document describes a PoC implementation that shows low end-to-end latency validated with real-time market data from NASDAQ using trusted measurements from industry-leading instrumentation offered by TS-Associates. The integration and access to the market is provided by Universal E-Business Solutions. This section briefly describes the value of the ecosystem partners in this PoC.

Cisco Systems 3548 Switch

The Cisco Nexus 3548 is a full-featured, ultra-low-latency cut-through Ethernet switch. The Cisco Nexus 3548 platform also brings industry-leading innovation in analytics with active buffer monitoring as part of the Cisco Nexus 3500 platform's Algorithm Boost (Algo Boost) engine. The enhanced feature set of the Cisco® Algo Boost engine also includes NAT, Switched Port Analyzer (SPAN), and robust Layer 3 capabilities. http://www.cisco.com/go/nexus3000.

The Cisco Nexus 3548 is designed to address next generation high performance trading market requirements enabling flexibility in deployment with three different deployment modes as shown below in Figure 2.

Figure 2: Forwarding Modes of Nexus 3548



Normal Mode: This is the default mode of the switch with full feature capability. Latency as low as 250 nanoseconds is possible. This mode provides maximum IP and MAC tables along with the entire enterprise features covering ECMP, QoS, NAT as well as full L3 routing. The latency is consistent regardless of what features are enabled. This is due to the fact that the Cisco Nexus 3548 architecture allows the packet parsing in parallel for all the features simultaneously without incurring latency hit in subsequent feature add.

Warp Mode: The Warp mode bypasses some of the normal-mode functions in such a way that the switch can parse the layer 2 and layer 3 packets in parallel. The Warp mode allows 20% lower latency compared to normal mode with targeted latency of 190 nanoseconds. Further the PoC validation below shows that latency as low as 140 nanoseconds is achievable. This mode is designed for various design requirements such as layer-2 only deployment or co-location switching where reducing latency more important than enabling all the features available in normal mode, some of which are restricted in Warp mode, with reduced but practical table sizes in most deployments.

Warp SPAN Mode: Warp SPAN allows packet forwarding with latency as low as 50 nanoseconds. It achieves such a low latency through bypassing table lookup and packet parsing. In this mode the specific port (1/36) acts as a source port and rest of the ports (47) are divide in groups of four as SPAN destination ports. With this mode any packet that is received on port 1/36 is replicated at the clock speed. Multiple groups of destination ports can be enabled at the same time. Many innovative designs and trading infrastructure enhancements are possible with this capability, such as optimized feed handler distribution, normalized data rebroadcast without the need of multicast as well as record-keeping and compliance with monitoring, This mode also can work in conjunction with Warp mode.

Warp mode and Warp SPAN can be enabled simultaneously and work together to enhance the trading topology with unparalleled reduction in latency as well as reduced cost in deploying multiple devices. The Warp mode essentially gives the same feature and functionality as normal mode for most trading environment and thus combination of normal mode and Warp SPAN is not required.

Note: The Warp SPAN will be supported after the first release of Nexus 3548 Nx-OS release. Please refer to http://www.cisco.com/go/nexus3000 for latest updates and requirements.

Enyx FPGA

Enyx is a leading developer of proprietary ultra-low latency technology, focused in delivering FPGA enabled applications for the financial industry. Enyx offers an ultra-low-latency FPGA-based market data filtering and distribution platform based on the high end Altera Stratix V board.

What is Feed Handling?

Exchanges broadcast market data using Multicast channels and UDP transport protocol. Each receiver needs to subscribe to one or more multicast channels to get the data corresponding to the traded symbols. Because of the nature of UDP transport protocol, where packets may be lost during transport, in order to secure the data exchanges broadcast the same data on two different multicast channels (also called line).

The trading system will subscribe to redundant lines A and B and, using sequence numbers, will detect packets lost and use the second line to recover missing data. This operation is called Feed Arbitration. Once feed arbitration is done, a feed handling system will filter out all unneeded symbols to keep only what the trading system needs. On US Equities market data feeds, this operation is required to keep track in memory of all live orders on the market and to know their associated symbols. This operation is called Order Management.

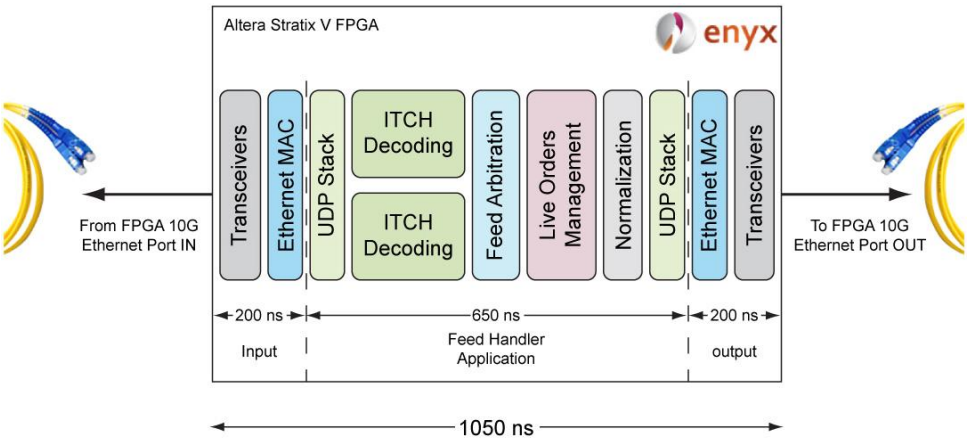
Output of the feed handling system is a Normalized data format representing the orders or the books. This data may be sent back to the network to reach multiple trading servers or directly through PCI-Express to the trading server memory using DMA (Direct Memory Access) technology.

Usually these operations are done using software, running on standard hardware. But the data volume keeps increasing year after year, and servers cannot handle processing data at 10Gb/s, resulting in queuing, packet drops, and latency increases up to several milliseconds during market data bursts. Therefore trading systems consuming more than one feed from multiple exchanges will need to deploy multiple Feed Handling servers to manage a fixed number of symbols with a decent performance.

Full FPGA Feed Handling

The block diagram below shows the FPGA application processing stages. The latency domain is split in two major processing blocks: the ingress/egress of data through the Ethernet medium; and the application block (consisting of UDP stack, decoding of specific exchanges' format, A/B arbitration, order management and normalization).

Figure 3: FPGA Feed Handler Application Block Diagram



The latency of the application block is directly correlated to the size of packet being processed and is traditionally optimized per market feed characteristics. The complete operation will be performed **in 1.2 µs on average**, and bellow 1.4 µs for 95 % of the packets' range. The system is resilient to network bursts, and will always process the data at the same speed, without queuing data or losing packets. This enables the possibility to process more than one market on a single FPGA board. The output format of the Feed Handler is NXFeed, a normalized binary data format enabling use of a single API for all exchanges. Feed arbitration, order management, symbol filtering and normalization are processed at 9.8 Gbps FPGA wire-rate access. The sample monitoring GUI is shown below depicting FPGA activities in detail, including packet rate, each block's statistics and market data relevant statistics. Notice the arbiter lane B is shown red as it was not used in this PoC.

Figure 4: FPGA Monitoring GUI

TS-Associates

TS-Associates TipOff is a precision instrumentation appliance designed for monitoring latency-sensitive electronic trading systems. When used for real-time latency monitoring, TipOff provides comprehensive support for market data and transaction flows with latency root-cause analysis and a variety of alerting, reporting, and third-party systems integration options. With a wide range of decoders for standards-based, vendor-middleware, and exchange-proprietary protocols, TipOff supports requirements across buy-side, sell-side, and processing-venue use cases. Equally at home in production systems monitoring and development performance optimization settings, TipOff can be used through all phases of the trading systems development and deployment lifecycle. A suite of licensable TipOff modules enables a TipOff deployment to be tailored to a customer's needs, according to the customer's use case, protocol decoding, and analysis requirements.

Universal E-Business Solutions and NASDAQ Live Data

Universal is at the forefront of the development and commercial deployment of transformative industry-changing technologies. Universal E-Business Solutions has teamed with industry leaders Cisco and Enyx to develop a holistic solution that empowers the financial services industry to achieve exceptional advantages with their high-bandwidth, ultra-low-latency initiatives. Universal's hosting services and content-based approach helps clients make better decisions, develop more effective initiatives, and more effectively measure success, resulting in a greater return on investment (ROI).

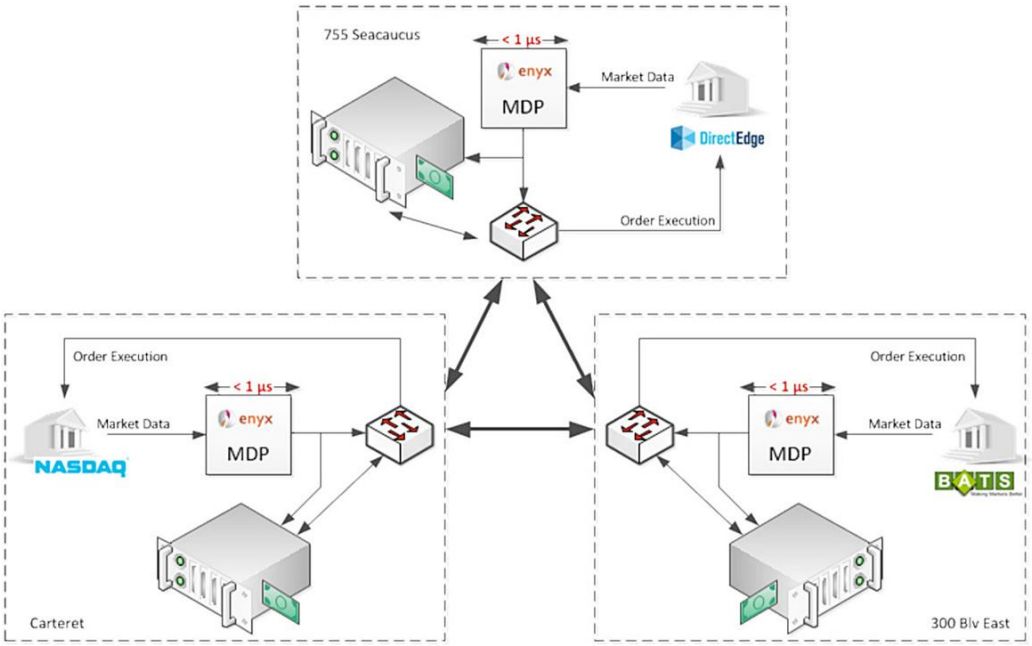
The latency-optimized PoC consists of:

- Cisco Nexus 3500 switch
- Enyx FPGA Market Data Processing card
- Data center hosting services and low-latency network
- Direct market data feed from the world's main exchanges

Trading companies often use the best available components from different specialist vendors. Selecting these components, integrating them, commissioning them, and deploying them for the best performance is challenging. To succeed in getting disparate components from heterogeneous vendors to work in concert requires an understanding of complex infrastructures plus the skills and experience that exceed the expertise of most small and midsize companies. As a comprehensive solutions integrator, Universal has painstakingly selected and tested the components in its latency-optimized turnkey solution to help ensure that they meet stringent requirements and exceed your expectations.

In this PoC a single data center with NASDAQ Feed was processed with a Cisco Nexus 3548 low-latency switch, and a trading server to hold the FPGA card and the trading system. Future PoC may show the total solution requires supporting multiple co-location and feed-procession required to support diverse market requirements.

Figure 5: Universal E-Business Solution Typical Setup with Three Data Centers

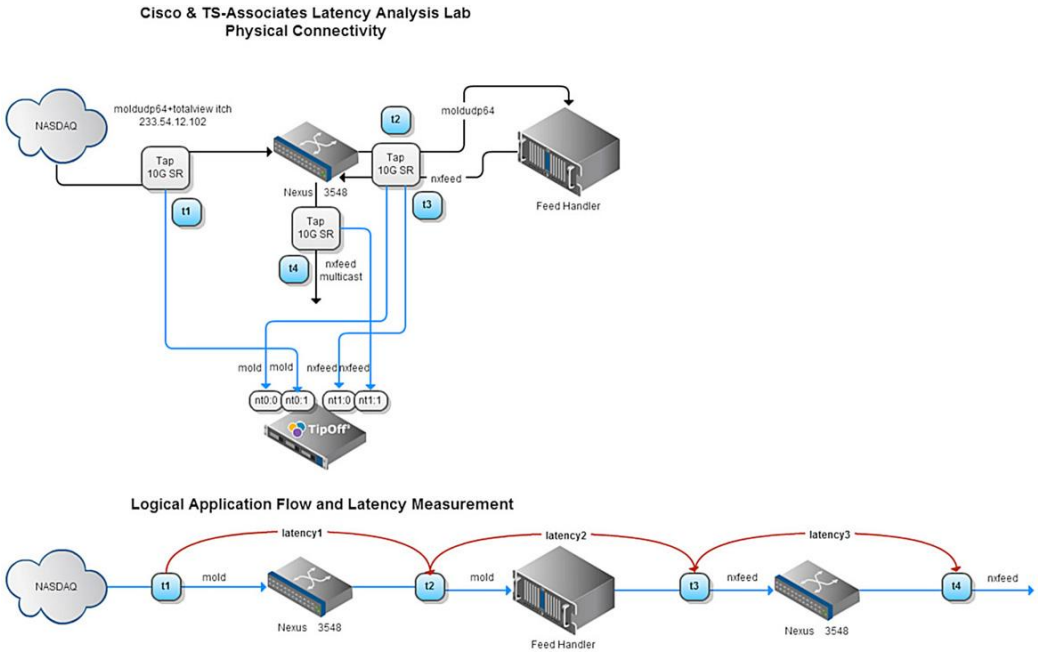


The use of an FPGA card for MDP (Market Data Processing) dramatically reduces the overall latency of the feed handling system and eliminates the need for an additional extra server in each location, reducing the system footprint and cost.

Validated Proof of Latency

In this validation, the Cisco Nexus 3548 Switch is commissioned at NASDAQ data center under the Universal E-Business Solutions license. Figure 6 shows the feed-handling systems and end-to-end latency hop.

Figure 6: End-to-End Latency Measurement



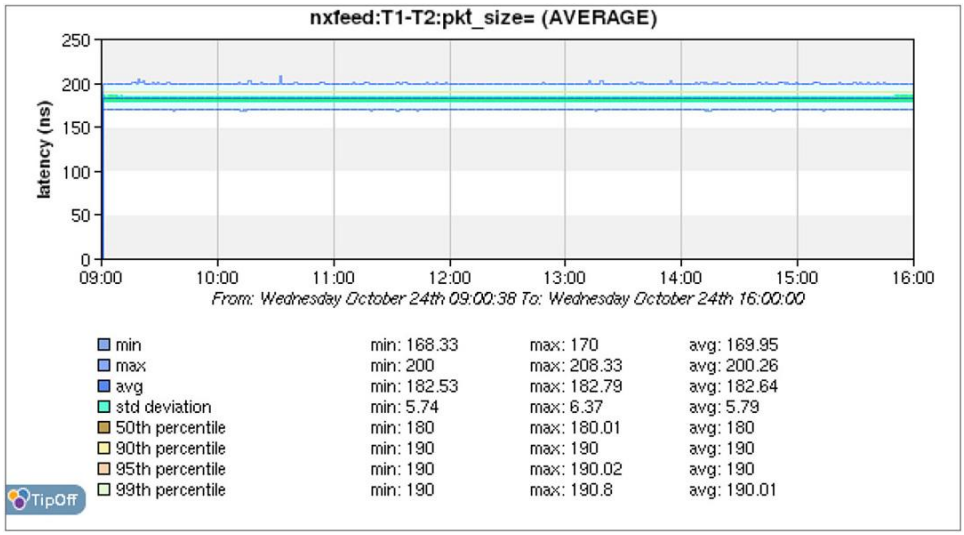
The NASDAQ ITCH (registered trademark) multicast feed is delivered through Cisco Nexus 3548 switch with multiple IXIA port providing an IGMP receiver support. The end-to-end per hop latency is monitored and measured with nanosecond precision using the TS-Associates TipOff server through Datacom Systems 10G passive tap. The PoC latency measurements are show below for two modes of Cisco Nexus 3548 (Warp and Warp SPAN). Each mode the first hop latency (T1-T2) is shown. The FPGA latency hop(T2-T3) remains unaffected by mode of operation and is thus only shown in Warp mode section.

Warp Mode Latency: As described in previous section Warp mode allows the bypassing lookup latency allowing achieving latency of 190 nanosecond or better. The following command enables the warp mode in Cisco Nexus 3548 switch:

```
NJX4-HOST-NX3548(config)# hardware profile forwarding-mode warp

The following command displays the state of the forwarding mode:

NJX4-HOST-NX3548# show hardware profile forwarding-mode
==========================
 forwarding-mode : warp
==========================
 host size = 8192
 unicast size = 4096
 multicast size = 8192
 l2 size = 8190
```

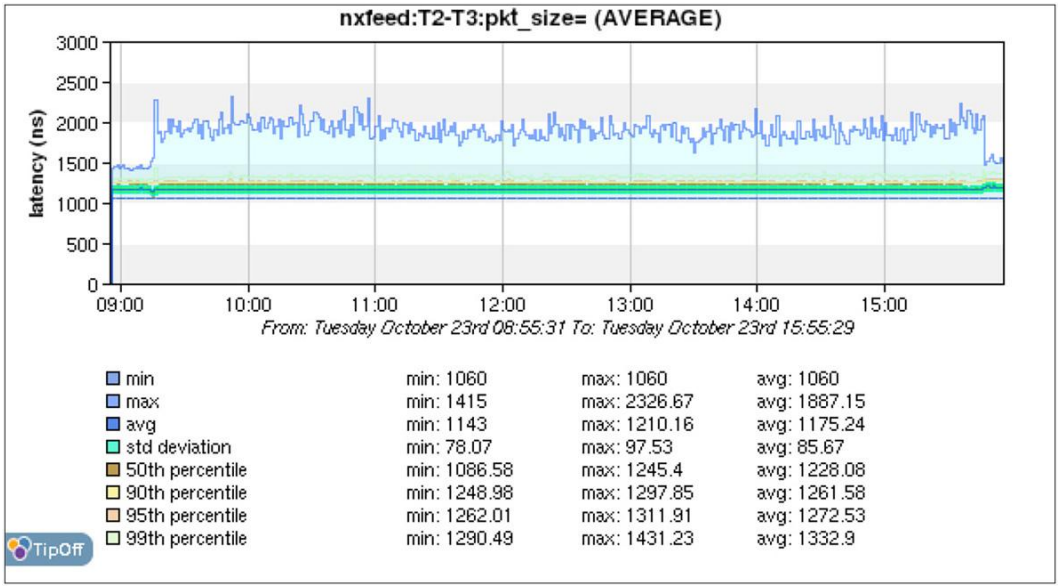
Figure 7 shows the latency for multicast feed traversing the first hop (T1-T2) depicted in Figure 6. As seen from the graph, the latency is well below 190 nanoseconds. In addition, the critical measurement that any high performance trading customer looks for is the latency for the 99th percentile. The 99th percentile indicates the spread of the latency, which can interfere with the trading strategy. Typically the lower the number, better the prediction of trading execution guarantees. Additionally the standard, deviation which provides the measure of how linearly the data is being transported across the infrastructure, is well below 10 nanoseconds.

Figure 7: TipOff Latency Measurement - Warp Mode Nexus 3548 - First Hop (T1-T2)



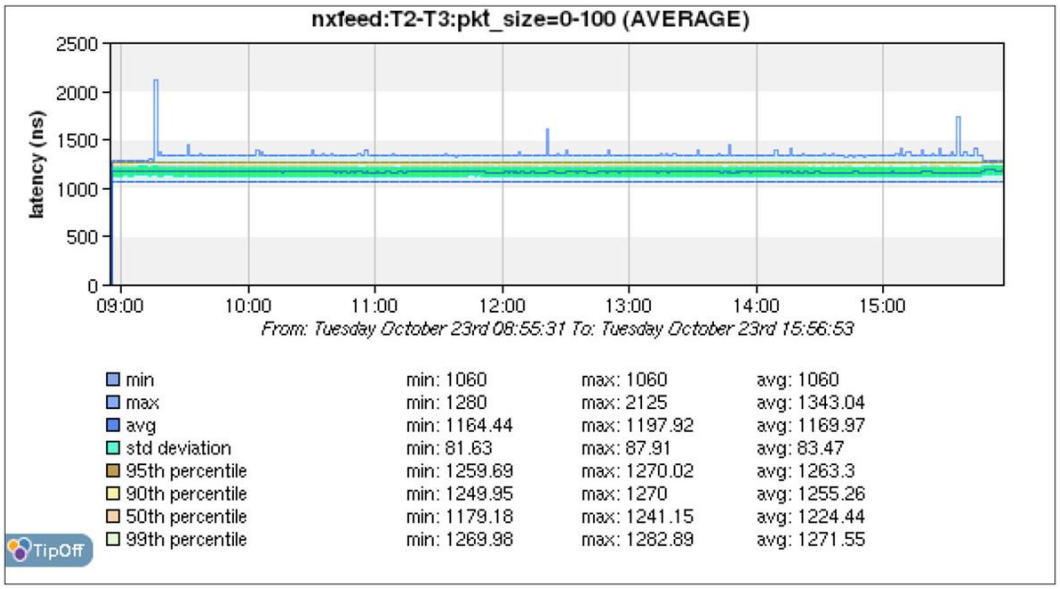
The below graph illustrates the latency in/out through FPGA for the hop T2-T3 as depicted in figure 6. In this live market data snapshot, FPGA processing time for data acquisition, processing and rebroadcasting shows minimum latency (1050 ns) and maximum latency (3080 ns). Maximum latency is seen for 1500 Bytes packets, at market opening after the auction period is over.

Figure 8: TipOff Latency Measurement Through FPGA Hop (T2-T3)



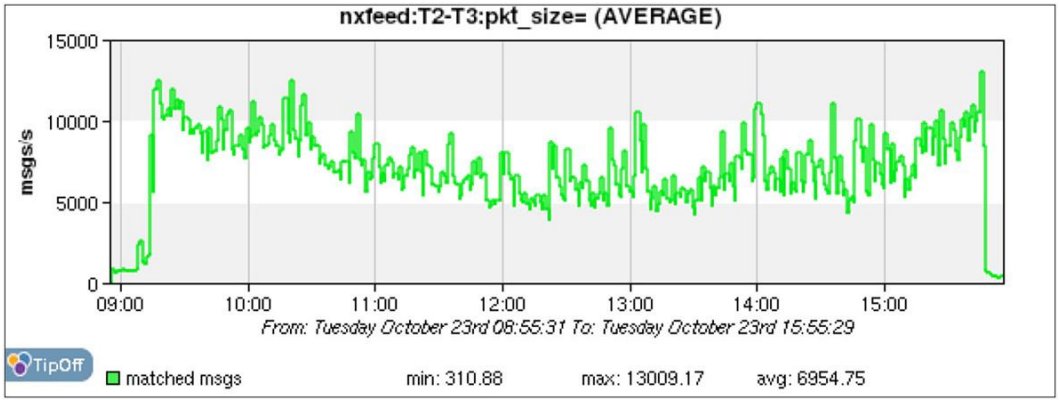
The graph below shows feed handler latency stability for 0 to 100 Bytes input messages. 0 to 100 Bytes messages represent more than 90% of the packets processed for the NASDAQ ITCH market feed. 99 % of the packets are processed in less than 1300 ns.

Figure 9: TipOff Latency - Packet size 0 -100 Byte



This graph below shows the bandwidth increase after the market opening at 09:30. The FPGA Feed Handling application is handling between 6000 and 14000 messages per second.

Figure 10: TipOff NASDAQ ITCH Message Counts



Warp SPAN*

As described in previous section Warp SPAN allows the latency associated with packet processing to be bypassed all together, achieving latency of 50 nanoseconds or better. Following command enables the Warp SPAN in Cisco Nexus 3548 switch:

```
NJX4-HOST-NX3548(config)# hardware profile warp span group 5

Second steps is to enable the destination port receiving the data via "switchport
monitor" command

NJX4-HOST-NX3548(config)# int e 1/17
NJX4-HOST-NX3548(config)# switchport monitor

NJX4-HOST-NX3500#show hardware profile warp span group all
```

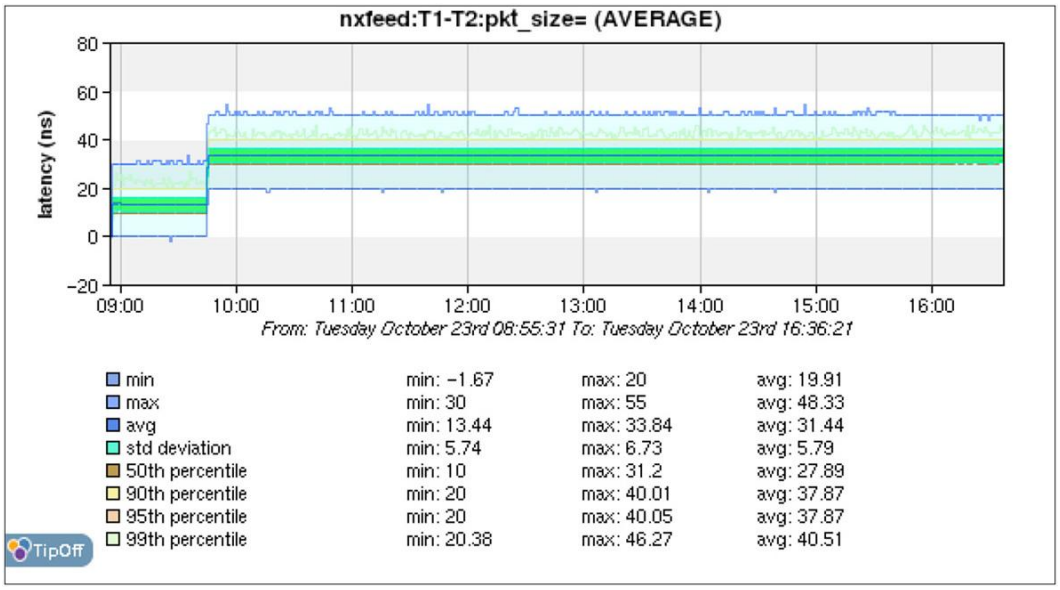
```
WARP SPAN GROUP:
Group Num ---- Port ----
 2 ( 5 --- 8)
 5 ( 17 --- 20)

 Total 2 warp span groups!
```

The first step is to enable the Warp SPAN for the group(s) of ports of that need to receive data in Warp SPAN. The data must be received on port 1/36 to be replicated at wire-rate. In this PoC the market data feed is received on port 1/36 and the switched at Warp SPAN speed to FGPA connected to port 1/17.

The below latency graph data is indicative of the Warp SPAN in action. As seen from the graph the average latency is well below 50 nanoseconds. Notice, also the 99th percentile latency is also below 50 nanoseconds.

Note: The TipOff sever has resolution in the range of +/- 10 nanoseconds, clock accuracy of +/- 0.005% and stamping accuracy of +/- 30 nanoseconds. The other variance such as cable-length and pre-FCS Nx-OS affecting measurement applies as well.

Figure 11: TipOff Latency Measurement - Warp SPAN Nexus 3548 First Hop (T1-T2)



Note: The Warp SPAN will be supported after the first release of Nexus 3548 Nx-OS release. Please refer to http://www.cisco.com/go/nexus3000 for latest updates and requirements.

The FPGA processing time remains the same with or without the Warp SPAN.

Summary

Cisco Nexus 3548 provides the lowest latency at line rate without loss of any practical features capabilities. The ENXY FPGA provides consistent latency, the lowest among available options for Feed Handler processing. The monitoring provided by TS-Associate at nanosecond level is crucial to guarantee the trust in vendors' claims to latency. This live market data PoC is a proof of the fact that the next generation infrastructure requires best of breed components, multi-vendor integration, market data access and integration expertise, enabled via firms like Universal E-Business.

About Cisco

Cisco Systems, Inc. designs, manufactures, and sells Internet protocol (IP)-based networking and other products related to the communications and information technology (IT) industry and provide services associated with these products and their use. The Company provides a line of products for transporting data, voice, and video within buildings. Its products are designed to transform how people connect, communicate, and collaborate. Its products are installed at enterprise businesses, public institutions, telecommunications companies, commercial businesses, and personal residences.

About Enyx

Enyx is a leading developer and provider of proprietary ultra-low latency technology focused on financial markets industry. Enyx's core value is to design and deliver the next generation of cutting-edge trading technology, while ensuring flexibility and high availability.

Enyx offers off-the-shelf solutions compatible with major FPGA platforms vendors, as well as assisting in the integration and deployment into customer infrastructure. Currently, these solutions are focused on market data acquisition from multiple and concurrent heterogeneous sources, normalization of the feed before its distribution, and pre trade risk. Its rapidly growing client base includes Stock Exchanges, ECNs, ATSs, proprietary trading firms, Investment Banks, Brokers and Investment Funds.

http://www.enyxfpga.com or contact@enyxfpga.com.

About Universal E-Business Solutions

Universal is a major technology provider in the high-performance computing area. Our architectural approach combines various cutting-edge technologies to deliver high-performance solutions for superior performance in a wide range of applications, including financial services, high-performance computing (HPC), cloud computing, and low-latency network and data centers. Universal focuses on transforming our client's IT environment to better serve their customers. Our goal is the alignment of people, process, and technology with business strategy. Universal offers several services to its financial trading vertical, including low-latency networks, managed LAN and WAN, network operations center monitoring, unified communications solutions and data center planning, and design and implementation services. Our solutions and products are used globally by many of the world's largest financial trading firms. Universal is headquartered in Hoboken, New Jersey. For more information about Universal products and services, visit http://www.uebiz.com or contact sales@uebiz.com.

About TS-Associates

TS-Associates began in 1999 as a consulting services firm, specializing in financial middleware products: deployment, integration, and custom development. Today, we are a hybrid product development and consulting services company. For more information, visit http://www.ts-a.com/products.html.

About Datacom Systems

Datacom Systems is a leading manufacturer of Network Test Access Points (TAPs), Data Aggregation Tools and other network access devices. Since our founding in 1992, we've built a reputation for quality engineering and unmatched customer service. Always an innovator, Datacom Systems is credited with inventing the revolutionary physical layer matrix switching technology. http://www.datacomsystems.com/index.asp.

C11-720080-00 10/12