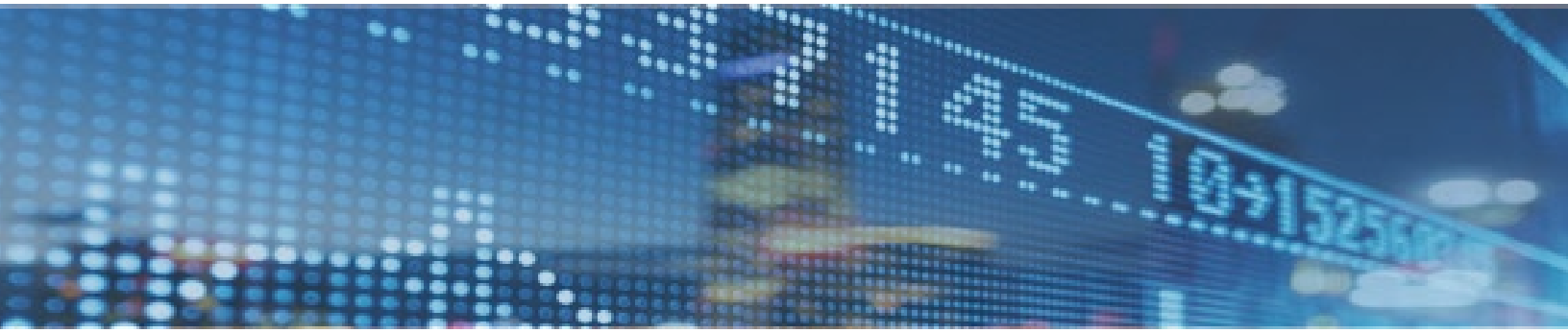
解决方案简介



借助新一代智能互连突破低延迟交易阻碍

搭载 Cumulus Linux 的 Mellanox SN2000 系列以太网交换机

10GBE 和更快速度上的最低延迟

随着算法交易的出现，对低延迟交易平台的需求变得非常迫切。在这些市场数据交换环境中，例如，在执行套利策略时，失去每一毫秒可能导致大约 100 万美元的机会丧失。因此，速度成为关键区分因素，尽可能降低处理环境之间的延迟至关重要。支持低延迟交易的公司有着极大的竞争优势，能够比竞争对手更快地存取流动性储备资金，在活动密集的高峰期内具有强于对手的表现。低延迟交易公司需要敏捷和高性能的基础架构，以处理通常与市场波动时期相关联的海量和激增的高速率数据。

基于 Mellanox Spectrum™ 的 SN2000 系列交换机搭载 Cumulus Linux，是业内最佳解决方案，能够满足低延迟交易公司的苛刻运营需要。该联合解决方案包括：

硬件支持：
- 从 10GbE 到 100GbE 系统的最低延迟直通式交换
- 16MB 共享缓冲区，用来处理突发流量

软件支持：
- 对低延迟流量的稳健、可扩展的协议支持（使用 PIM、SM 和 SSM 的多播）
- 使用直方图和基于水印的触发器加强缓冲区监视

统一运营模式：
- 增强的命令行支持（采用 NCLU，即 Cumulus Linux 网络命令行实用程序）
- 与 Linux 系统工具/应用程序的原生集成

自动化：将 Linux 服务器生态系统扩展到交换机上
- 对所有自动化系统的原生支持，包括 Ansible、Puppet 和 Chef

亮点

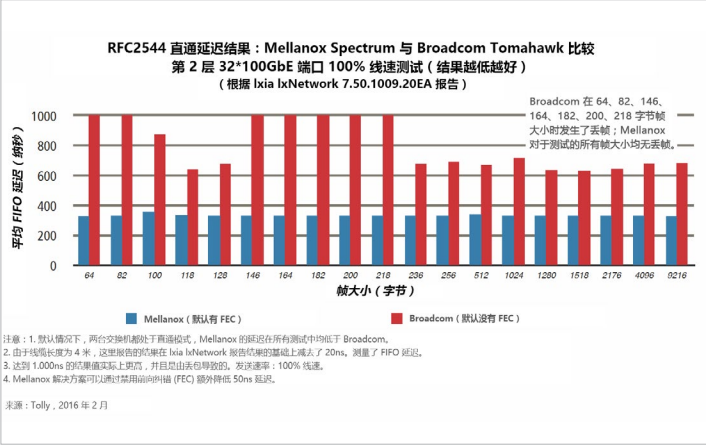
- 延迟更低，可达 300ns
- 16MB 共享缓冲区，用来处理突发流量
- 实时缓冲区监控
- 高级流量镜像
- 大规模多播
- 零丢包 www.zeropacketloss.com
- IEEE 1588 PTP
- 自动化就绪平台

金融信息交换的消息传递性质形成可能压制网络元素的突发流量。Cumulus Networks 和 Mellanox technologies 密切协作，克服突发流量的影响。Spectrum 卓越的共享缓冲区设计，具有可预测的线速性能和零丢包，以及对突发流量的无可匹敌的适应性，从而实现绝对的延迟公平性。Cumulus Networks 还在缓冲区配置和监控方面达到前所未有的灵活性，确保消费者可以得到完全的系统功能。

每一纳秒都很珍贵……

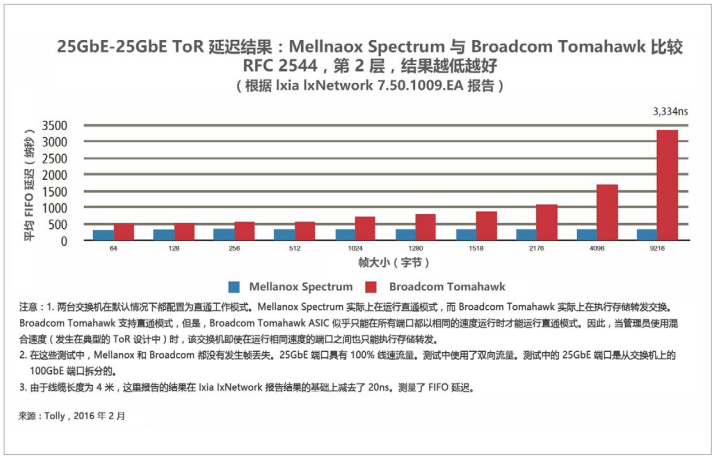
交换机有一个任务，就是尽可能快地转移每一帧。丢弃帧和/或引入额外延迟可能对应用产生重大影响，因为对于电子执行服务的需求和提供，微秒级和纳秒级体现的是一种竞争优势。

基于 Mellanox Spectrum 的平台在所有帧大小和所有链路速度下都展示了可持续的 300ns 直通式延迟和零丢帧。而且，即使在兼有更慢和更快速度端口的混合环境下，仍可保持此延迟。



构建数据中心的基础前提之一是网络基础架构的执行方式需要可预测。可预测性的衡量标准是吞吐量和延迟的一致性，无论包大小，无论需要时在用户之间划分流量有多公平。对于低延迟交易，不公平分配可能会导致租户的性能较差，并使其丧失预测和控制流量行为的能力。

构建和维护同步、准确的时序解决方案的能力是成功预配置和管理高频率交易 (HFT) 网络和应用程序的基础。使用 IEEE 1588 PTP，交换机可以为现有网络基础架构内的应用程序提供高度准确的精确时间同步，而无需投资和部署单独的定时网络。Spectrum 和 Cumulus Linux 低延迟解决方案与 Cumulus Linux 软件栈相配合，为 PTP 提供支持。



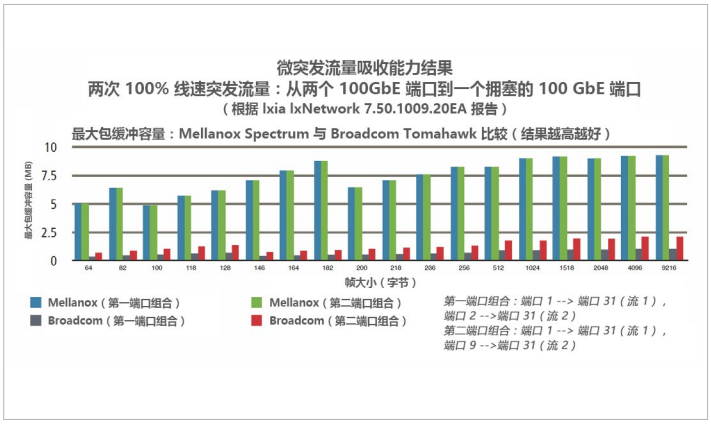
即使在突发流量期间性能也可预测

突发流量是导致网络出现短暂拥塞的流量模式。这通常源自导致网络端点向网络中发送突发流量的高度活跃时段，例如，在市场的高波动期间。即使超出网络带宽的突发流量导致一个信息包丢弃，也可能大幅增加交易时间，需要重新传输，并可能导致整个数据丢失，从而导致交易中断。

由于市场数据的突发性质以及吸收和汇总各种来源的需求，每次重传都可能导致大量收入损失。面临的挑战是最大限度地降低端到端的交易延迟，同时处理峰值而不丢包。

突发流量吸收测试用于测量在等待输出端口变得可用期间，交换机可在缓冲区中保存的帧数。缓冲区越大，丢弃的流量越少，从而避免性能下降。

您可以清楚地看到，端口组合之间的差异极小，清楚地展示了该体系架构的好处。



增强的缓冲区监控和可见性

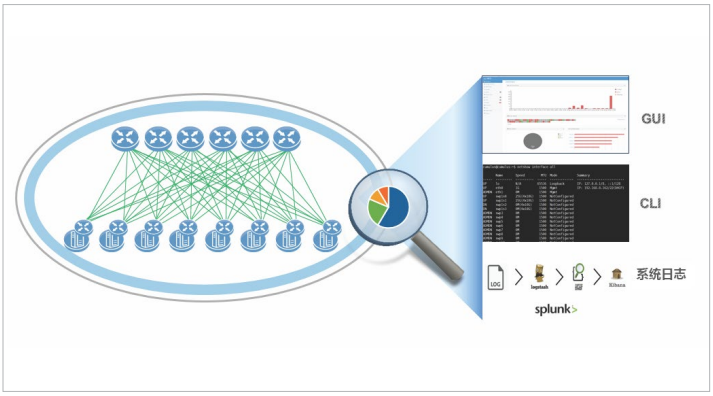
微突发流量会导致非优化网络不堪重负，并导致事务中断或需要重传，这会对公司的财务业绩产生显著的负面影响。利用新型 100GbE 交换机，1 秒可以转换成超过一亿条消息！

拥塞监控事件不会坐等网络拥塞的影响向上渗透到应用层，它可以在应用层受到影响之前主动检测即将发生的拥塞事件，以便抢先进行容量规划。

Spectrum 提供实时硬件缓冲区监控。该体系架构让交换机在缓冲数据包时，能够检测在通信流中发生的延迟。检测在硬件级别是否存在拥塞有助于了解交换机内部的情况。了解拥塞的位置、深度和持续时间有助于进行额外优化。

能够对每个交换机端口配置警报阈值，只要离该端口而去的流量进站或出站队列造成流量延迟数毫秒以上，便会创建一个系统日志条目。利用记录的系统日志，交易商能够分析网络在高流量/拥塞事件期间的历史性能表现，在必要情况下分配额外带宽。

Cumulus Linux 在该组功能之上构建，使 Linux 应用程序能够注册和监控缓冲区深度、高水位阈值跨越和完全计数器补充。这表现为一个流式的异步接口，以 JSON 格式提供，以便更高级别的工具使用。这种组合提供了一个简单、快速、强大的方法，能够监控、分析网络流量，准确定位网络中发生的流量问题。



提供、保持和真正证明遍历交换机时的延迟的能力对于好的高频率交易解决方案是不可或缺的。如以上所示，Spectrum 与 Cumulus Linux 的组合为可预测、不变的延迟、有保证的公平性和吞吐量提供最佳解决方案。

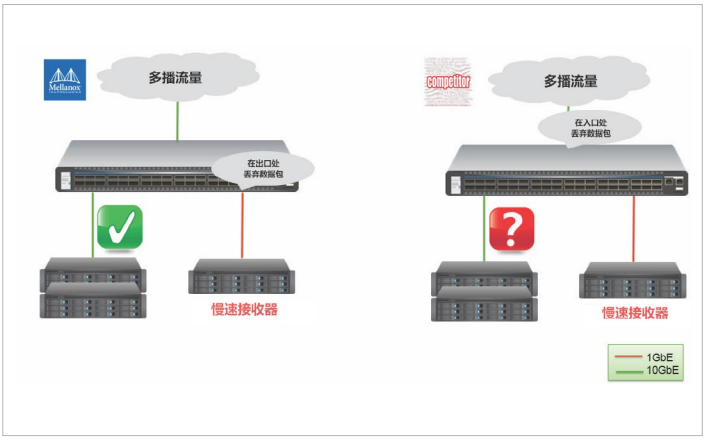
此外，借助创新的缓冲区监视技术、缓冲区控制工具和高质量 PTP 实现，该解决方案有能力验证这些功能，并对实际流量模式和行为表现提供深刻见解。

多播提升...

IP 多播是一种常用技术，专为"少对多"IP 数据传输而设计，并在金融市场数据馈送网络中广泛使用。通过物理连接，交易商能够接收市场数据（通常通过 TCP 或 UDP 多播）和发送订单（通常通过 TCP 或 UDP 单播）。低延迟交易平台使交易商能够在几秒钟内执行数百万笔订单并扫描多个市场和交易所。因此，低延迟交易商严重依赖大规模多播流量。

Spectrum 交换机采用 Cumulus Linux 软件栈，完全支持多播协议，包括协议无关的多播稀疏模式 (PIM-SM)、PIM 指定源多播 (PIM-SSM) 和多播源发现协议 (MSDP)。Cumulus Linux 支持大规模多播，与市场中可用的任何其他解决方案相比，支持的 MAC、IPv4 路由和多播组数量更多。所有这些都通过了广泛的验证和互操作性测试，以确保从当前的 1/10GbE 基础架构快速迁移到下一代 25/100GbE 部署。

在 Cumulus Linux 多播堆栈支持之上，Spectrum 针对多播线头 (HOL) 阻塞进行了优化，在 HOL 阻塞中，单个慢速接收器就可能导致交换机发生拥塞（当馈送推送速度超过该接收器的吸收能力时）。对于运行多播流量的系统，在入口处执行丢弃，交换机也会"惩罚"快速服务器，因为多播数据包在复制之前即被丢弃。另一方面，Spectrum 交换机仅在慢速接收器出口端口执行丢弃，从而使快速服务器（多播组的成员）不受影响地继续运行。



总结

Mellanox 的 Spectrum 与 Cumulus Linux 相结合，代表市场上在 10Gbps 和更快速度上实现最低延迟和抖动的唯一解决方案。该联合解决方案在延迟、抖动、突发流量吸收和低功耗等方面体现了 Spectrum 的最佳性能，并具有丰富的第三层软件栈和 Cumulus Linux 的监控功能。非常适合需要低延迟和高频多播大规模套件的客户。

规范

交换机型号	SN2700	SN2410	SN2100
100GbE 端口的最大数量	32	8	16
40GbE 端口的最大数量	32	8	16
25GbE 端口的最大数量	64	64	64
10GbE 端口的最大数量	64	64	64
吞吐量	3.2 Tb/s	2 Tb/s	1.6 Tb/s
每秒包数	4.77 Bpps	2.98 Bpps	2.38 Bpps
延迟	300ns	300ns	300ns
CPU	双核 x86	双核 x86	ATOM x86
系统内存	8 GB	8 GB	8 GB
SSD 内存	32 GB	32 GB	16 GB
数据包缓存	16 MB	16 MB	16 MB
100/100 管理端口	1	1	1
串行端口	1 个 RJ45	1 个 RJ45	1 个 RJ45
USB 端口	1	1	1 个迷你 USB
热插拔电源	2（1+1 冗余）	2（1+1 冗余）	无
热插拔风扇	4（N+1 冗余）	4（N+1 冗余）	无
气流换向选件	有	有	有
典型功耗 (ATIS)	150W	165W	94W
尺寸（宽度 x 高度 x 深度）	1.72″ x 16.84″ x 27″ （43.9 毫米 x 427.8 毫米 x 686 毫米）	1.72″ x 17.24″ x 17″ （43.9 毫米 x 438 毫米 x 436 毫米）	1.72″ x 7.87″ x 20″ （43.9 毫米 x 200 毫米 x 508 毫米）

关于 Mellanox

Mellanox Technologies 是针对服务器和存储的端到端 InfiniBand 及以太网互连解决方案和服务的领先提供商。Mellanox 互连解决方案可提供最高吞吐量和最低延迟，更快地向应用程序传递数据并充分发挥系统性能，从而提高数据中心效率。Mellanox 提供一系列快速互连产品：适配器、交换机、软件、线缆和芯片，它们可针对广泛的市场（包括高性能计算、企业数据中心、Web 2.0、云、存储和金融服务）加快应用程序运行时间并最大程度实现业务成果。

深入了解 Mellanox 产品和解决方案：

www.mellanox.com

关于 Cumulus Networks

Cumulus Linux 具体实现了原生 Linux 网络连接。该系统提供的超级加强版内核以及其他与网络连接相关的软件包，运用了网络连接方面最新的行业理念，同时保留了与 Debian 提供的全系列软件的兼容性。运行 Cumulus Linux 的 SN2000 系列提供标准网络连接功能，例如桥接、路由、VLAN、MLAG、IPv4/IPv6、OSPF/BGP、访问控制、VRF 和 VXLAN 叠加 (Overlay)。

深入了解 Cumulus Networks 操作系统：

www.cumulusnetworks.com

北京市朝阳区望京东园七区保利国际广场 T1 15 层
Tel: 010-5789 2000
www.mellanox.com