

使用 Mellanox 高效网络助力下一代 IP 广播

目录	IP 网络面临的广播挑战 :	2
	Mellanox 面向下一代广播基础架构的解决方案	3
	结论	7

简介

下一代基于 IP 的高性能演播室目前正在进行广播行业改革，尤其是在广播制作领域。这是因为如今使用的串行数字接口 (SDI) 不仅效率低下，而且复杂且有定制性质。迁移到基于 IP 的基础架构让广播公司能够在内容创建和分发、多平台支持及未来视频格式的所有领域进行创新。随着开放式网络和商业现货 (COTS) 解决方案的出现，助力当今数据中心和云的同一技术为此创新提供了底层基础。

随着向基于全 IP 的演播室转移越发稳健，大型演播室正在考虑部署在 IP 而不是 SDI 上。先进媒体工作流程协会 (AMWA) 和联网媒体联合工作小组 (JT-NM) 领导的跨多个组织联合协作项目已经加快了基于 IP 演播室的创新和采用。这为电影电视工程师协会 (SMPTE)、视频服务论坛 (VSF) 和音频工程协会 (AES) 等组织的多个新标准铺平了道路，如：

- 将 SDI 整体包装在 IP 数据包中 (SMPTE 2022-6)
- 发现和注册协议 (AMWA IS-04)
- 以多个 IP 流的形式分离出实体流¹ (TR03 (ST2110), TR04)
- 管理成功的时间同步传输 (PTP – RFC1588, SMPTE2059)

但是，就像任何技术创新一样，这些新标准带来了不同程度的复杂性，广播工程师需要加以理解并采用。

¹ 在广播领域，所有媒体内容都由两个部分组成：实体和元数据。实体数据是实际的视频和音频媒体。此数据可以是一个或多个的各种实体数据类型，包括压缩或未压缩的视频以及任何数字音频源

IP 网络面临的 广播挑战：

数据包 (IP) 与电路 (SDI) 交换解决方案之间有着本质的区别。因此，通过此范式转移交付解决方案意味着，需要在 IP 网络上小心地采用广播 workflow。以往，IP 网络架构设计为携带 IT 流量，这与广播媒体的需求不同。服务质量 (QoS)、缓冲区和暂停帧全都会给网络延迟带来变化。带有 PFC (优先级流量控制) 的 QoS 和暂停帧设计为可解决争用，从而允许相应地设置流的优先级，而缓冲区则在网络拥塞时保存数据包。

SDI 是一个基于电路的交换解决方案，因此在 SDI 中交换源是即时瞬间发生的。但是，新一代的视频格式 (如 4K 和 8K) 以及诸如高帧速率 (HFR) 和高动态范围 (HDR) 等广播要求需要在无拥塞的网络上采用一种新的体系架构。当视频流 (IP 数据包) 穿过 IP 架构时，它们与具有相同优先级的其他流穿插在一起。在一个专门建造的 IP 工厂中，可以控制环境以确保管理拥塞并保证视频交付。在更为公开共享的网络中，例如在大型体育赛事中使用的网络，控制谁可以访问哪些资源以及什么时候访问可能具有一定的挑战。

要想成功地传输音频和视频，需要具有在基于 IT 的传统环境中看不到而且很大程度上也不需要的可预测性、可靠性和确定性。规范使用 IP 资源的方式；通过限制多个同步发送人的速率来减轻对交换机缓冲区的影响；确保在架构上均匀地分配流量都是 IP 广播工厂的一些技术建议。

数据包需要均匀地分散或分发到可用网络资源上，以确保正确地分发负载，从而解决资源争用和公平性问题。由于视频可用的变量过多，例如压缩、未压缩、HD-SDI、3G-SDI、4K、8K、HFR 和 HDR，因此高带宽和一系列的端口速度成为基本要求。扩建一个能够向所有观众公平分配流量的无阻塞可伸缩架构是能否成功采用 IP 广播基础架构的关键。

Mellanox 面向 下一代广播基础 架构的解决方案

在过去的两年间，Mellanox 通过 JT-NM、AMWA 和 SMPTE 互操作性测试和技术讨论一路领先，并且一直与 Fox、BBC 和 NBC 等重要广播公司展开合作，帮助定义和交付下一代 IP 演播室。

Mellanox Technologies 是一个领先的端到端以太网和智能互连解决方案供应商。Mellanox 行业领先的技术可提供最高吞吐量和最低延迟、更快地向端点传递数据并充分发挥系统性能，从而提高 IP 演播室的效率。Mellanox 提供了一系列的高性能解决方案：网络和多核处理器、网络适配器、交换机、线缆、软件和芯片，可以加快并平衡吞吐量并获得最大成效。

并不是所有技术都是一样的。Mellanox 的技术可提供基于 IP 的端到端演播室，如下所示：

Mellanox 低延迟交换机

Mellanox 基于相同的芯片提供许多固定外形规格的交换平台，因此各平台之间的性能已知，且已经过测试和证明。

- 多种带宽连接 (1G、10G、25G、40G、50G 和 100G) :
支持各种视频要求所需的吞吐量，包括 4K、8K、HFR 和 HDR (不管有没有压缩)。25Gb 以太网可向客户端交付 4K 无压缩流，而 100GbE 用于高速交换机互连。
- 可预测的缓冲区分配：
灵活的大型缓冲池可用于交换机上所有端口，而不是划分为各个端口组。灵活的零丢包交换机缓冲区能够存储广播应用中普遍存在的突发流量，并确保根据需要转发所有数据包。

- QoS 和 DSCP 标记：
允许标记重要的流，以辅助确保架构无阻塞。
- 可预测的网络性能：
上一代基于 SwitchX-2 以及当前一代基于 Spectrum 的交换机都始终如一地为所有数据包大小交付可预测的性能。这在 Tolly 测试报告中进行了介绍和很好的记录。

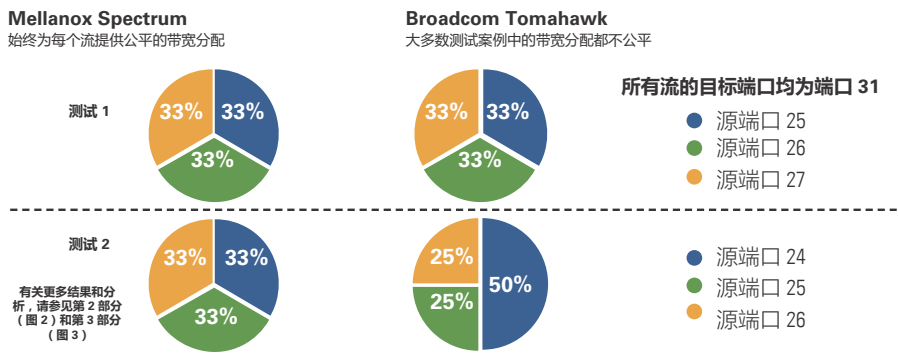


图 1. Mellanox Spectrum 的可预测网络

- 一致且非常低的端口到端口延迟和抖动：
Mellanox 交换机具有行业最低的端口到端口延迟和抖动。与其他 IP 交换机供应商一起，此交换机也经过了 FOX 网络工程和运作的测试，并且得到了 Snell Advanced Media (S-A-M) 的确认。Fox 根据这些测试总结得出，经证明，Mellanox 平台是交换机芯片性能方面当之无愧的领导者。端口到端口延迟以及数据包延迟变化 (PDV) 也经过了测试，在所有接受测试的供应商中，经证明，Mellanox 具有最低且最确定的延迟。当基于 IP 的广播演播室使用 Mellanox 架构进行扩展时，它使用大家熟知的已知常量进行该操作。早先的时候，Fox 和 Aperi Corporation 已经在 2014 年度 SMPTE 技术会议展览会上在其他交换机供应商之间呈现过测试结果。

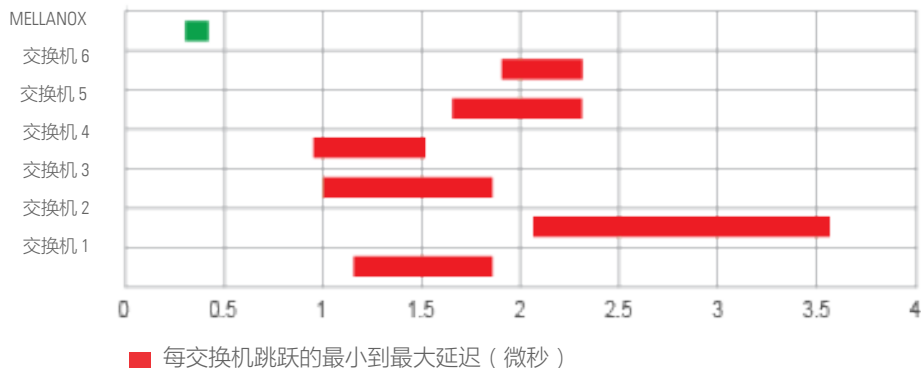


图 2. 端口到端口 (交换机到交换机) 延迟

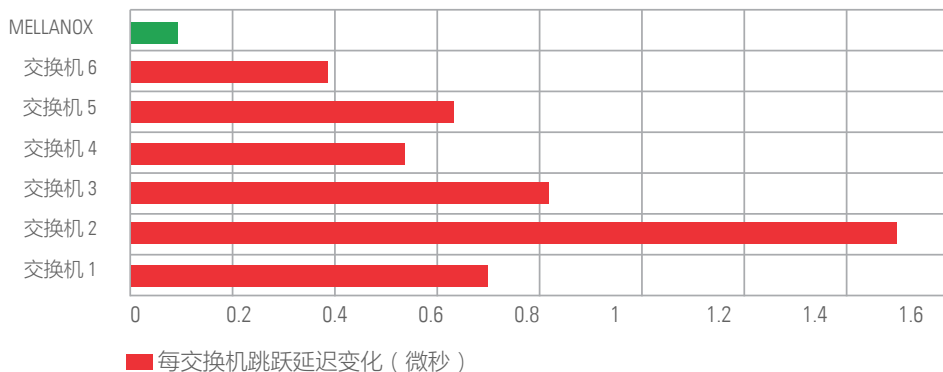


图 3. 数据包延迟变化 (抖动)

• 架构容器化广播服务：

通过将 IP 演播室服务容器化并在交换机上运行它们，广播工程师们便可以将精力集中在为他们的工作室打造理想 IP 媒体架构上，无需使用额外的服务器和虚拟机。例如，Mellanox 可使用在交换机中运行的 NMOS 注册和发现服务及 NMOS 感知连接管理器来构建 NMOS 感知交换机。

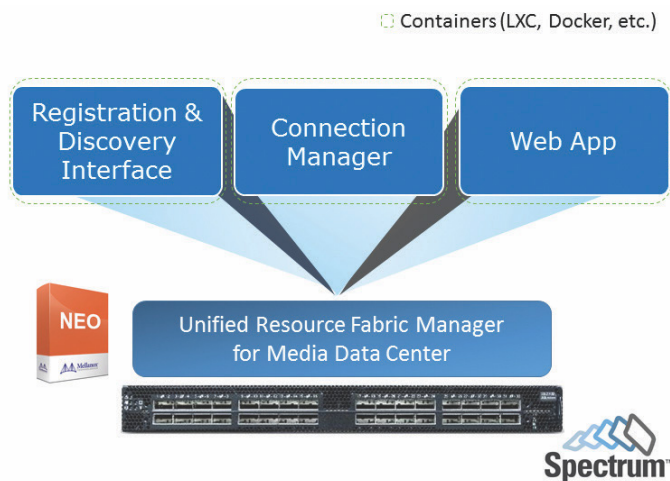


图 4. 在 Spectrum 交换机上容器化 IP 演播室服务

• OpenFlow 集成：

使用 OpenFlow，Spectrum 交换机允许无缝配置多播到单播和单播到多播转换、流复制以及基于 NAT 的流交换。Spectrum 支持最新的 OpenFlow 1.3 标准，它允许在整个架构内操纵交换机内的流。使用其 RESTful API，OpenFlow 可动态控制在整个网络架构内转发广播流的方式。而且，Spectrum 允许支持多个配置：仅 OpenFlow、基于端口混合和基于协议的传统操作模式。它也与 Open Daylight (ODL)、开放网络操作系统 (ONOS) 和其他控制器相集成。

Mellanox 网络适配器

网络接口适配器 (NIC) 在将可预测的高性能广播流交付到网络方面发挥着重要的作用。

- 多种带宽连接：1、10、25、40、50、100GbE
- 内核旁路：

Mellanox 基于 ConnectX 的网卡支持多个内核旁路解决方案，包括 RoCE（基于融合以太网的 RDMA）、Netmap、VMA 和 DPDK（数据平面开发套件）。这提供了更快的数据包处理速度，从而实现更低的延迟和更高的吞吐量。内核旁路技术减少了 I/O 处理的上下文切换和内存复制开销。这样就给应用返回了更多的 CPU 周期，从而降低抖动并提高吞吐量。

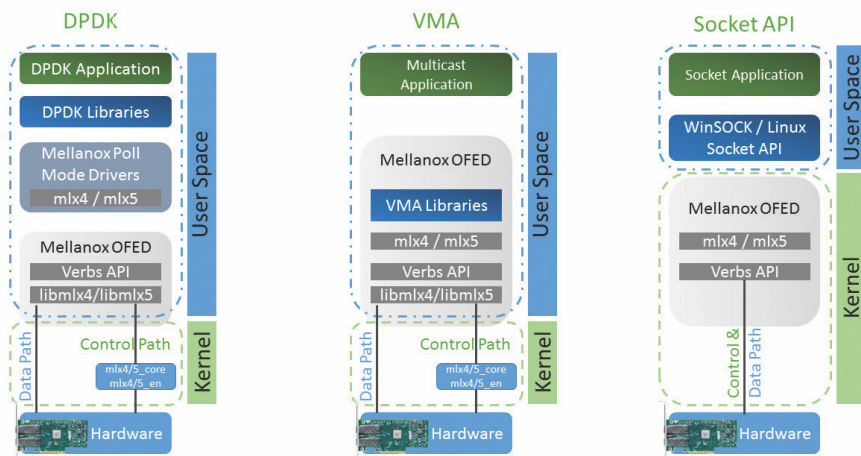


图 5. Mellanox 内核旁路解决方案

Netmap：来自用户空间的线速原始数据包 I/O，比萨大学开发

Mellanox VMA 消息传递加速器：一个动态链接的用户空间 Linux 库，用于加速基于 UDP 的多播流量，以便从 CPU 卸载网络处理。绕过内核和 IP 堆栈将最大程度减少上下文切换、缓冲区复制和中断，从而使延迟变得极低。

数据平面开发套件 (DPDK)：用于进行快速数据包处理的一组数据平面库和网络接口控制器驱动程序

通过在适配器内使用如上所述的实现功能，我们就可以选择让哪些数据使用内核堆栈，让哪些绕过内核。

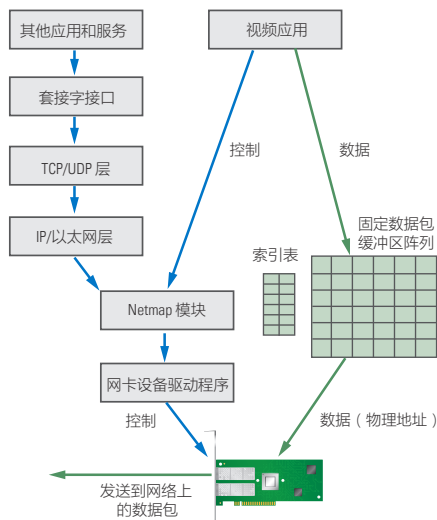


图 6. 与 BBC 联合开发的优化 Netmap 解决方案

- 使用数据包同步防止网络拥塞：数据包同步克服了多个同步流全都同时发送数据，从而导致交换机缓冲区冲突和溢出的问题。设想一个具有大量突发发送人的 IP 网络，当发往网络的视频流同步时，突发数据包数量也会同步，这可能会导致端口由于缓冲区耗尽而阻塞。在无阻塞架构中，需要在服务器和交换机中同时解决和消除该问题。通过让服务器上的所有发送人放慢流出速度，这个问题就可以得到缓解。

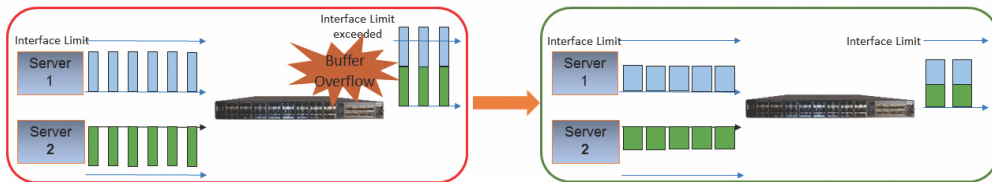


图 7. 使用数据包同步防止网络拥塞

- 硬件时间戳：基于 IEEE1588 时间同步的精确度取决于可用时钟的精确度以及在哪里终止合适的时间戳。这些时间戳将依次用于计算时钟与真实时间的偏差/偏移。Mellanox 硬件时间戳克服了服务器硬件上由软件堆栈引入的差异。通过在硬件网卡终止时间戳，Mellanox 适配器就可以消除在基于软件的解决方案中看到的高度不可预测的抖动。

Mellanox 光缆和铜缆

通过任何架构传输视频都要求具有非常高的性能和精确度。所有 Mellanox 互连都按照非常高的标准进行构建，可支持和维护我们高速度和低延迟的公司口号。不管在 SFP 还是 QSFP 外形规格中，无源铜缆和有源光缆（VCSEL [垂直腔面发射激光器] 硅光）及光收发器全都以最低电源和最低成本为理念进行构建。经测试达到行业最低 10e-15 的比特误码率 (BER)，这意味着与竞争产品相比，它们具有更少的传输错误和重试次数。

结论

由于专有 SDI 架构效率低下，向下一代基于 IP 的广播世界前进已不可避免。随着像 BBC 和 FOX 这样的公司在其基础架构中使用 IP 网络，这一趋势甚至变得更加明显。Mellanox 端到端高效广播网络迅速改变了广播应用的经济情况，包括广播专业人士的现场体育报道、制作演播室、数据传输、内容分发和存储应用。使用已经过证实的可伸缩解决方案，Mellanox Spectrum 交换机、ConnectX 系列适配器和 LinkX 线缆让广播公司能够节省时间和资金，同时将极其可靠的 HDR 视频交付给其观众。

Mellanox 是 IEEE、SMPTE、AMWA 和 JT-NM 标准的积极参与者，帮助进一步推动了媒体和娱乐行业所需的 IP 网络要求。

词汇表

1. AMWA – 先进媒体工作流程协会 - <http://www.amwa.tv>
2. JT-NM – 联网媒体联合工作小组 - <http://www.jt-nm.org>
3. SMPTE – 电影电视工程师协会 - <https://www.smpte.org>
4. NMOS
5. VSF – 视频服务论坛 - <http://www.videoservicesforum.org>
6. AES – 先进工程协会 - <http://www.aes.org>
7. VCSEL – 垂直腔面发射激光器



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