

vivo 与英特尔优化 HDSL B 方案 显著改善负载均衡系统性能

“蓬勃发展的互联网业务带来了网络负载压力的不断增长，凸显了负载均衡系统性能优化的必要性。通过 HDSL B 方案、E810 网卡的采用，以及我们与英特尔在转发性能优化、吞吐性能优化等方面的合作，我们能够在 TCP、UDP 等广泛的场景中实现性能的显著提升，这有助于我们更好地控制 TCO，推动业务的敏捷发展。”

— 侯畅

vivo 互联网基础运营研发总监

概述

如今，网络业务正在实现蓬勃发展，需要连接的用户、设备、应用数量正在快速增长，这在带来了更加丰富多彩的网络服务、提升数据价值的同时，也给网络基础设施的性能、稳定性等带来了巨大的挑战。通过在现有的网络结构之上建立负载均衡系统，企业能够利用透明故障转移机制、性能监控等方式，将流量引导到最佳可用服务器，从而提供一种廉价、有效、透明的方法扩展网络设备和服务器的带宽，增加吞吐量，加强网络数据处理能力，提高网络的灵活性和可用性。

作为中国知名的智能终端和智慧服务品牌，vivo 依托规模庞大的网络基础设施，为海量用户提供了高效、高稳定的网络服务，显著提升了用户体验。vivo 利用针对英特尔® 硬件平台定制优化的高密度可扩展负载均衡器（High Density Scalable Load Balancer，HDSL B）软件，对负载均衡方案进行了深度优化；同时，vivo 还采用了英特尔® 以太网网络适配器 E810，满足了高吞吐量业务的需求，改善了网络应用的性能。

挑战：爆发式数据流量给负载均衡系统带来巨大压力

在互联网业务中，访问流量常会由于用户访问的时间、空间差异而发生大幅波动，如果不能及时对这些流量进行引导与处理，很可能导致用户访问缓慢乃至停止响应，给用户体验带来极大的负面影响。作为在线业务高并发、高可用系统必不可少的关键组件，负载均衡应运而生，其目标是尽力将网络流量平均分发到多个服务器上，以提高系统整体的响应速度和可用性。

凭借着手机等移动终端业务的快速发展，以及持续的业务创新，vivo 已经成为全球具备重要影响力的移动互联网服务提供商。数据显示，截至 2022 年 5 月，vivo 服务的国内移动互联网用户就已经超过 2.7 亿，vivo 的应用商店服务于 2.5 亿月活用户，其中日活用户超过 7056 万，已经累计实现超过 1.1 万亿的应用分发总量，商店的日应用分发量达到 10.8 亿次¹。而庞大的用户数也意味着，互联网业务已经成为 vivo 重要的业务增长点。

¹ 数据援引自 vivo 内部数据。英特尔并不控制或审计第三方数据。请您审查该内容，咨询其他来源，并确认提及数据是否准确。

互联网业务量级的增长使 vivo 负载均衡系统承受着巨大的压力，这些压力主要来源于：

- vivo 高并发的业务数量正在快速增长，尤其是智能手机等终端数量的增长带来了持续的业务流量，给负载均衡系统带来了更高的要求。
- vivo 日益增长的网络带宽流量，尤其是大数据业务，对网络带宽提出了更高的要求。
- 在互联网交易等敏感业务场景中，对于业务实时响应速度有着更高的要求，vivo 希望通过更高性能的负载均衡系统来提升业务响应速度。

为了应对上述挑战，vivo 要求负载均衡系统的单机平均需要承载近百万的每秒查询率（QPS）请求，同时要达到 500W/s 以上的包处理能力。显然，一般的负载均衡服务器根本无法达到这么大量的请求和包处理速度。vivo 希望对当前的负载均衡系统进行进一步的优化，以满足业务需求

解决方案：基于 HDSL B 与英特尔® 以太网网络适配器 E810 的负载均衡方案

vivo 的负载均衡系统主要分为四层负载均衡与七层负载均衡。四层负载均衡为七层负载提供更上一级的负载均衡。在本白皮书中，vivo 与英特尔在负载均衡上的合作优化主要围绕着四层负载均衡而展开。

在四层负载均衡服务器中，vivo 使用了英特尔® 以太网网络适配器 E810。该网络适配器具备 100/25GbE 性能，支持单个或双端口连接，在 PCIe 4.0 x 16 插槽中提供了出色的性能，并支持应用程序设备队列（ADQ）、动态设备个性化（DDP）、RDMA iWARP 和 RoCEv2 等各种高级功能，能够有效满足各种工作负载对于带宽和时延的严苛要求。

在四层负载均衡中，vivo 还使用了基于英特尔® 平台的软硬结合四层负载均衡优化方案：HDSL B。HDSL B 是基于英特尔® 平台的软硬结合四层负载均衡优化方案。HDSL B 的第一期设计目标主要定位为解决性能挑战。其“高密度”是指单节点并发连接和数据吞吐的高密度，“可拓展”主要指性能可以随着 CPU 核数量或者资源总量的增加线性拓展。

对于单个物理服务器负载均衡节点，HDSL B 的初步设计性能目标包括：实现亿级最大并发连接数；单 CPU 核吞吐 8Mpps 以

上，且能够随 CPU 核数线性增长；TCP 连接新建速率单 CPU 核 80 万 CPS 以上，且能够随 CPU 核数线性增长；并发连接数量不随并发连接数增长而下降。

除了性能方面的特点之外，HDSL B 也具备一些一般负载均衡产品常见的产品特性，如高可用，支持多节点间的 Session 同步，在负载均衡模式方面支持 FNAT/DNAT/SNAT 和 DR 模式。负载均衡算法还支持常见的一些种类，如 RR、WLC 以及 Consistent Hash，同时也具有安全和可观察性方面的特性。为提升转发性能，英特尔提出了快速路径（Fast Path）与缓速路径（Slow Path）的方案来提升转发效率。

为提升吞吐量，英特尔协助 vivo 实现了英特尔® 以太网网络适配器 E810 100g 的网卡适配，并支持基于 IP 和全端口的 FDIR 功能来满足高并发量的要求。双方基于英特尔® 数据平面开发套件（英特尔® DPDK）20.08 版本进行优化，满足业务转发多队列场景一致性的需要。vivo 充分利用了 E810 的 ADQ 特性。该特性支持将隔离的硬件网卡队列供某些应用程序进行专有使用，并且使这些硬件网卡队列可以最佳地连接到所需应用程序中的不同线程，从而改善网络应用的性能。

验证：性能实现显著提升

为了验证在 TCP 与 UDP 场景中 HDSL B 方案相对于 DPVS 所能带来的性能提升，vivo 在业务环境中进行了测试，测试组网架构如图 1 所示，测试配置如表 1 所示。配置使用的负载均衡服务器采用了英特尔® 至强® 金牌 6230 处理器，网卡则采用了英特尔® 以太网网络适配器 E810。

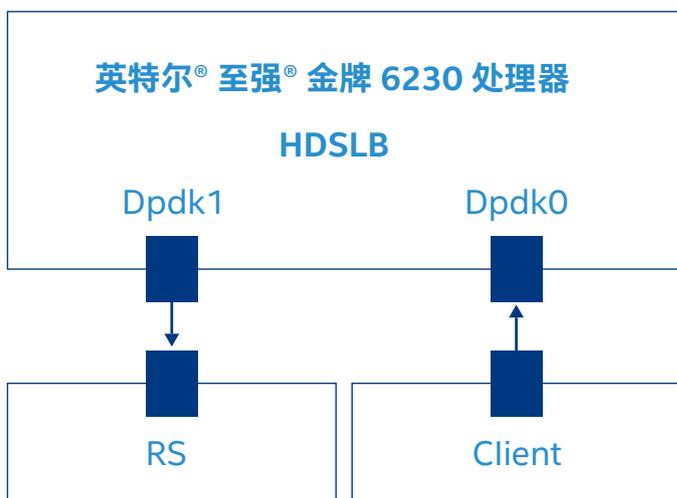


图 1. 测试组网架构

组件	描述
处理器	英特尔® 至强® 金牌 6230 处理器 @ 2.10GHz 内核数量: 80 性能测试使用 8 个 worker 内核
内存	251 GB
网卡	英特尔® 1592 (rev 02) (100 G)
操作系统	CentOS Linux release 7.4.1708 (Core)
Linux 内核版本	4.19.91
GCC 版本	gcc version 4.8.5 20150623 (Red Hat 4.8.5-39) (GCC)
DPDK 版本	20.08

表 1. 测试软硬件配置

在 TCP 场景中, vivo 采用了 wrk 来进行测试。wrk 是一款针对 HTTP 协议的基准测试工具,它能够在单机多核 CPU 的条件下,使用系统自带的高性能 I/O 机制,通过多线程和事件模式,对目标机器产生大量的负载。该场景的测试结果如图 2 所示,在单核情况下, HDSL B 方案相比 DPVS 方案,性能提升了 55%²。

在 UDP 场景中, vivo 采用了 hping 3 测试工具。hping 3 是面向命令行,用于生成和解析 TCP/IP 协议数据包的开源工具。测试结果如图 3 所示,在单核情况下, HDSL B 方案相比 DPVS 方案,性能提升了超过 200%⁴。

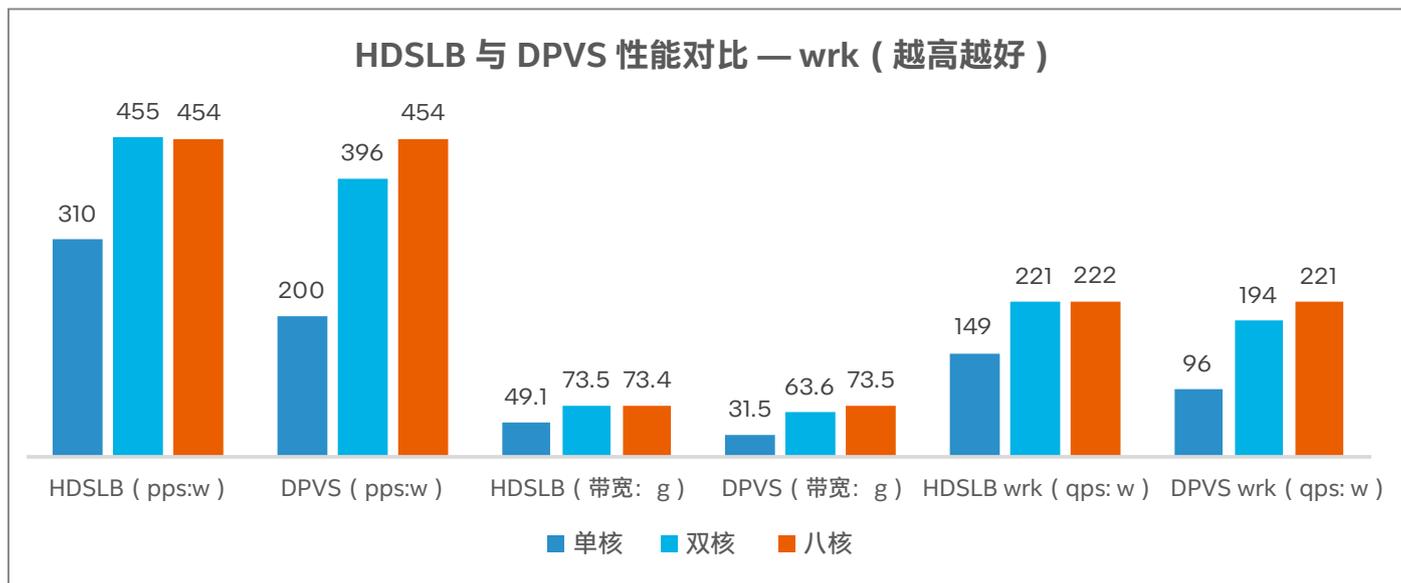


图 2. TCP 场景测试数据³

测试数据显示, HDSL B 方案在 TCP 与 UDP 场景中实现了较大的性能提升,能够提升四层负载均衡系统应对网络压力的能力。此外,受益于单节点性能的巨大提升, vivo 能够减少需要部署的负载均衡服务器数量,从而降低总体拥有成本 (TCO)。

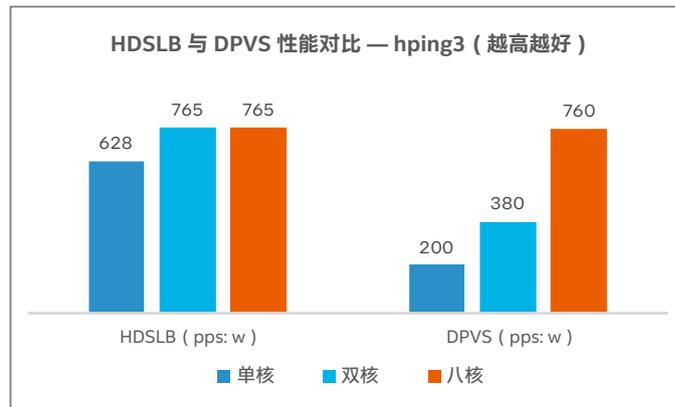


图 3. UDP 场景测试数据⁵

展望: 在大象流/老鼠流检测算法与 HQoS 等方面深入合作

未来, vivo 与英特尔将致力于在负载均衡上进行进一步合作,这包括大象流/老鼠流检测算法和实现,以及基于 HQoS (分层 QoS) 的限速功能。

^{2,3,4,5} 数据援引自 vivo 内部测试结果。测试配置见表 1。英特尔并不控制或审计第三方数据。请您审查该内容,咨询其他来源,并确认提及数据是否准确。

大象流/老鼠流检测算法和实现：依据网络流携带数据量的多少可将其分为大象流和老鼠流。大象流占网络流总数的 10% 左右，但其承载了网络总数据量的 85%，具有传输时间长、占用带宽大、数量较少的特点。老鼠流则正好相反。传统流量调度方法对所有网络流实行无差别的调度，而大象流较多的网络节点负载较大，各节点不均匀的负载会降低网络吞吐量、增加网络响应时间。如能快速、准确识别网络中的大象流并实现有效的负载调度，能减小网络阻塞的概率，提高网络的可靠性。

基于 HQoS 的限速功能：传统 QoS 技术可以满足语音、视频以及数据等业务的不同服务需求，可以针对不同的业务提供不同的服务。但是随着网络设备的高速发展，接入用户数量和每个

用户的业务量不断增多，传统 QoS 无法做到同时对多个用户的多个业务进行流量管理和调度。HQoS 采用多级队列调度的方式为多用户、多业务提供精细化的 QoS 服务，并通过多级队列进一步细化区分业务流量，对多个用户、多种业务等传输对象进行统一管理和分层调度，在现有的硬件环境下使设备具备内部资源的控制策略，既能够为高级用户提供质量保证，又能够从整体上节约网络建设成本。

通过上述合作，vivo 将有望进一步提升负载均衡系统的性能，从而支撑更大规模的业务访问，为用户提供更高质量、响应更加及时的互联网服务。

关于 vivo

vivo 是一家以设计驱动创造伟大产品，打造以智能终端和智慧服务为核心的科技公司。vivo 致力于打造追求乐趣、充满活力、专业音质、极致影像、愉悦体验的智能产品，并将敢于追求极致、持续创造惊喜作为 vivo 的坚定追求。vivo 充分利用本地的人才资源，布局了全球化研发网络，覆盖 10 个城市，业务范围涵盖 5G 通信、人工智能、工业设计、影像技术等众多个人消费电子产品和服务的前沿领域。

关于英特尔

英特尔 (NASDAQ: INTC) 作为行业引领者，创造改变世界的技术，推动全球进步并让生活丰富多彩。在摩尔定律的启迪下，我们不断致力于推进半导体设计与制造，帮助我们的客户应对最重大的挑战。通过将智能融入云、网络、边缘和各种计算设备，我们释放数据潜能，助力商业和社会变得更美好。如需了解英特尔创新的更多信息，请访问英特尔中国新闻中心 newsroom.intel.cn 以及官方网站 intel.cn。



实际性能受使用情况、配置和其他因素的差异影响。更多信息请见 www.intel.com/PerformanceIndex

性能测试结果基于配置信息中显示的日期进行测试，且可能并未反映所有公开可用的安全更新。详情请参阅配置信息披露。没有任何产品或组件是绝对安全的。

具体成本和结果可能不同。

英特尔技术可能需要启用硬件、软件或激活服务。

英特尔未做出任何明示和默示的保证，包括但不限于，关于适销性、适合特定目的及不侵权的默示保证，以及在履约过程、交易过程或贸易惯例中引起的任何保证。

英特尔并不控制或审计第三方数据。请您审查该内容，咨询其他来源，并确认提及数据是否准确。

© 英特尔公司版权所有。英特尔、英特尔标识以及其他英特尔商标是英特尔公司或其子公司在美国和/或其他国家的商标。其他的名称和品牌可能是其他所有者的资产。