

使用适当的网络适配器优化您的云端

目录

当今的云	2
网络适配器对云的价值	2
最高的速率和灵活性	3
通过将网络任务卸载到网卡来增加虚拟机密度	3
面向未来的解决方案	3
使用 SR-IOV 的裸机式虚拟机 (VM) 性能	4
利用 RDMA 实现高存储性能和大幅 COGS 节省	5
高性能存储 iSER	5
使用 SMB Direct 的 RDMA	5
使用高级 Open vSwitch (OVS) 卸载提高 CPU 利用率	6
数据中心密度和异构性	6
易部署	7
用于领先云平台的 CloudX 参考体系架构	8
OpenStack 集成	8
VMware vCloud 集成	8
Windows Azure Pack 集成	8
Mellanox 适配器满足您的云需求	9

当今的云

随着按需计算的需求不断增长，数据量也呈指数增长，需要更大、更高性能和更高效的云端，无论是公共云还是私有云。

各种云工作负载和技术都在增加 CPU 利用率的负担。叠加 (Overlay) 网络协议、OVS 处理、使用各种新协议的大规模存储访问以及其他高性能工作负载需要大量处理。这会浪费成本高昂的 CPU 周期，使其无法用于网络，并最终使大量带宽未得到利用。最终结果是应用程序效率受限，整个云也变得低效。

由于面对这些挑战，数据中心管理员现在希望实施智能、灵活的网络，可以提供足够的带宽来满足应用程序和存储需求，减轻 CPU 负载，并使云能够高效地扩展。智能网络可以通过卸载尽可能多的网络任务来分担负载，从而释放 CPU 计算资源来为更多的用户提供服务并处理更多数据。

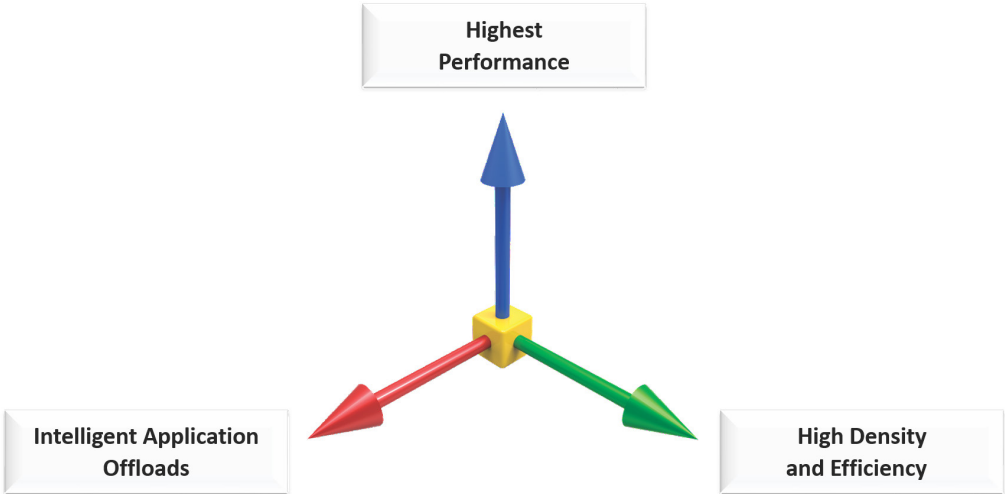


网络适配器对云的价值

网络适配器是高效的高性能云的基础。因此，在顶层运行的应用程序将能够以最小的限制进行扩展。

它们为云带来的优势主要体现在三个方面：

- **高性能：**云服务需求的增长需要增加带宽速率。
从一个端口提供 10Gbps 到 100Gbps 的带宽。向后和向前兼容性、按需增长的能力、将各种速率整合到一个数据中心的能力对于充分优化云效率至关重要。
- **网络智能：**当今数据中心设计的新趋势是在各个组件之间分担工作负载处理。这不再是仅限于 CPU 的任务，现代网络需要高级加速功能，包括从 RDMA（用于各种用例和应用程序）到协议特定的卸载（如 TCP、叠加 (Overlay) 网络和存储卸载）。在网络层上实施它们将释放昂贵的 CPU 周期以用于用户应用程序，同时提供改进的用户体验。
- **横向扩展效率：**与以往不同，现在的网络层对于云的高效扩展能力发挥着重要作用。网络加速有助于增加每台服务器承载的虚拟机数量，而其他新型创新技术（如多主机）可以支持构建最具成本效益、空间效率和节能的云。



最高的速率和灵活性

在云计算服务器上放置更多虚拟机，以及拥有服务更多用户和访问更多数据的应用程序，这些需要高达 100G 的带宽以及包含各种速率的灵活数据中心设计。

当今的云不再使用一种连接方式，而是根据计算节点和存储服务器的规模，需要一系列的解决方案。

Mellanox ConnectX-4 是市场中的首款 100G 适配器，也是首款 25G 和 50G 适配器。这些速率最适合新的流量密集型云和 Web 2.0 数据中心。

面向未来的解决方案

ConnectX-4 适配器系列使云可以根据需要从 10GbE 扩展到 25GbE，而无需更换硬件或软件。

速率/适配器系列	ConnectX-4	ConnectX-4 Lx
10 GbE	是	是
25 GbE	是	是
40 GbE	是	是
50 GbE	是	是
100 GbE	是	-

通过将网络任务卸载到网卡来增加虚拟机密度

智能网络将卸载网络和存储任务，这样做有两个主要优点：首先，释放 CPU 周期以供应更多虚拟机，其次，快速网络提供更多可用带宽。因此，可以有效地供应更多虚拟机。

其中一个例子是叠加 (Overlay) 网络协议实现，如 VXLAN、NVGRE 或 GENEVE。

在传统的实现中，当数据包被封装并且其格式发生变化时，适配器无法再对流动数据包执行网络卸载，从而将此工作留给 CPU。这导致大量 CPU 周期花费在网络任务而不是用户工作负载上。

这最终限制了总带宽以及可以放置在任何一台服务器上的最大虚拟机数量。

Mellanox 适配器可以解析和了解网络叠加 (Overlay) 协议 (VXLAN、NVGRE 和 GENEVE)，从而能够卸载其网络处理。这样可以实现更高带宽，大幅提高虚拟机密度，显著优化 CPU 利用率，从而带来巨大的成本效益。

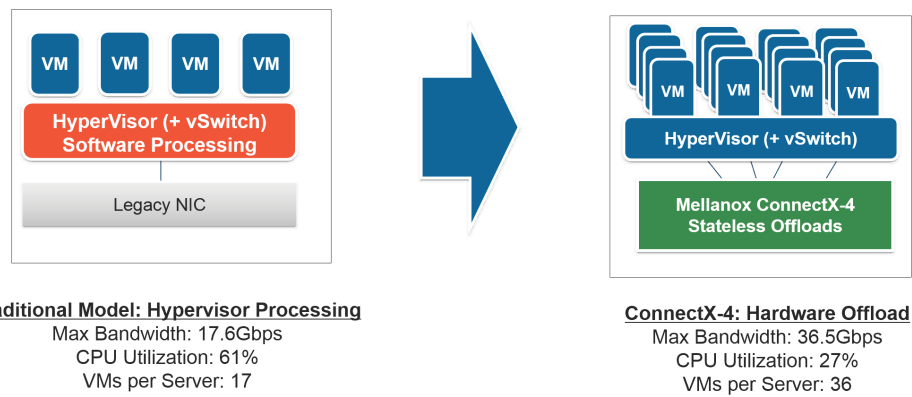


图 1. 具有 VXLAN 流量的 40GbE 适配器示例：虚拟机密度和带宽几乎翻倍，而 CPU 利用率降低了 35%

使用 SR-IOV 的裸机式虚拟机 (VM) 性能

Hypervisor 网络软件对虚拟机性能造成重大损失。因此，对于要求苛刻的高性能应用程序，云管理员传统上使用专用的裸机服务器，因为他们不认为虚拟机能够提供合适的网络性能。

但是，借助合适的网络技术（如 SR-IOV 和 OVS 卸载），现在可以从虚拟机运行需要低延迟和高带宽（例如金融和 HPC 工作负载）的应用程序。

Mellanox 通过启用 SR-IOV 实现虚拟机直接访问网络。Mellanox SR-IOV 技术还提供了从虚拟机到网络上的其他虚拟机或物理主机的 RDMA 访问，从而在虚拟机上实现了类似裸机的延迟和性能。

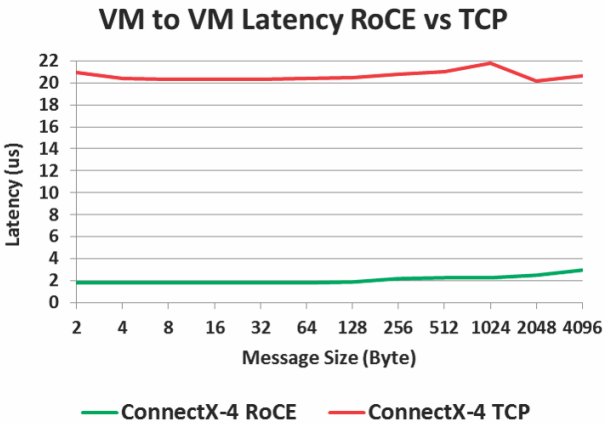


图 2. 基于 SR-IOV 的 RoCE 的延迟比 TCP SR-IOV 优秀 10 倍

利用 RDMA 实现高存储性能
和大幅 COGS 节省

通过在存储访问中启用 RDMA（远程直接内存访问）和 RoCE（基于融合以太网的 RDMA）连接，云用户现在可以受益于最低延迟和最高带宽，而不会浪费 CPU 上的任何额外负载。这样可以实现线速性能，同时节省计算和存储成本。

一些主要存储协议和平台现在都支持 RDMA。例如 iSER（基于 RoCE 的 iSCSI）、SMB Direct、OpenStack Cinder 中的 iSER、Ceph RDMA 以及其他协议和平台。

下面是 iSER 和 SMB Direct 的性能优势：

高性能存储 iSER

使用 iSER（用于 RDMA 的 iSCSI 扩展）进行存储通信可显著改进延迟、IoPS、带宽和 CPU 利用率。

在图 3 中可以看到使用带 iSER 而不是常规 iSCSI 的 Mellanox 适配器的优势。

iSER 现在嵌入多个存储解决方案和软件包中，例如 OpenStack Cinder。

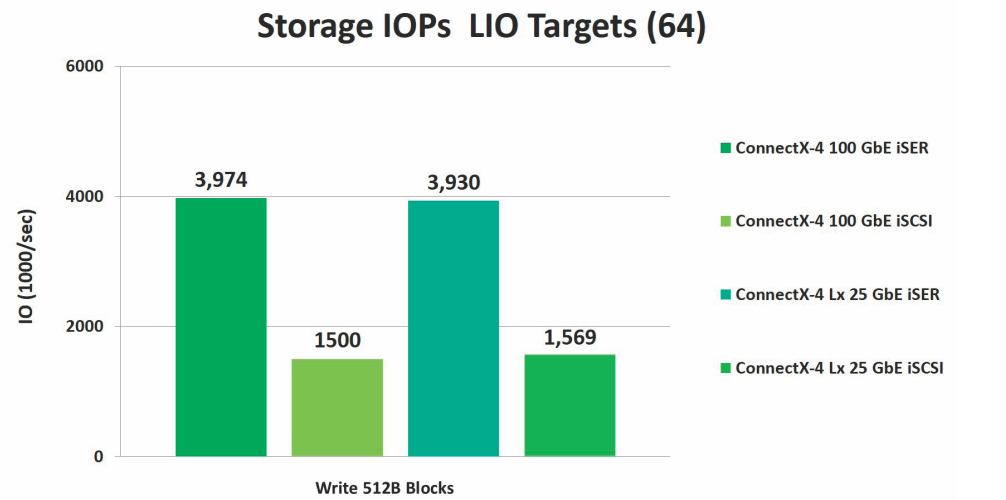


图 3. Mellanox iSER 与常规 iSCSI 的性能比较

使用 SMB Direct 的 RDMA

启用 RDMA 存储的另一个例子是 Microsoft SMBDirect。图 4 显示了使用基于 RoCE 的 SMB Direct 时获得的高性能。

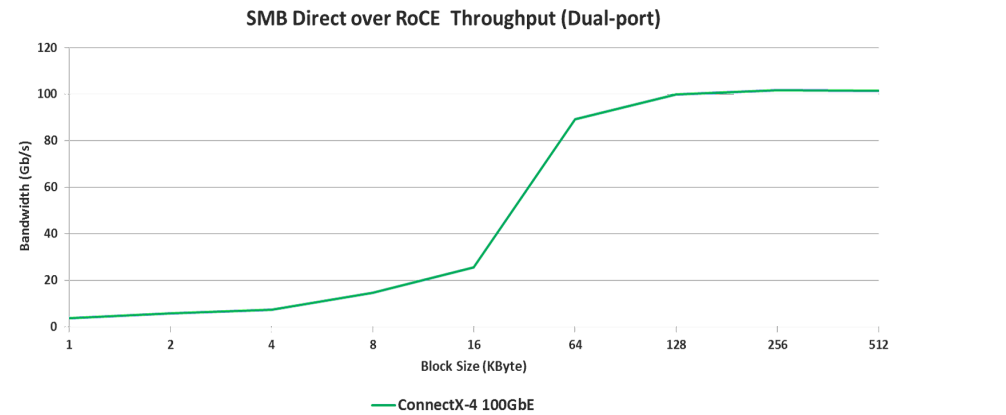


图 4. 使用基于 RoCE 的 SMB Direct 时的高性能

使用高级 Open vSwitch (OVS)
卸载提高 CPU 利用率

在软件中实施诸如 OVS 交换之类的网络任务会让主机 CPU 耗费在网络任务上，而不是专注于运行用户的应用程序。这导致巨大的成本负担和性能低效。

Mellanox ConnectX-4 可以实现虚拟机交换的硬件卸载，从而降低 CPU 开销并节省成本。

ConnectX-4 具有细粒度的按流卸载策略。可以针对每个流做出将其卸载或在 CPU 内处理的决定。如果在硬件中处理流量，则可以实施高级策略和卸载，例如叠加 (overlay) 网络协议卸载、访问控制列表、转发等。

OVS 卸载使用户能够在不改变应用程序行为的情况下继续使用 OVS，同时仍然可以从 SR-IOV 的高性能和 ConnectX-4 的硬件卸载中获益，从而提高带宽的可用性，实现高数据包速率，并改进 CPU 利用率。

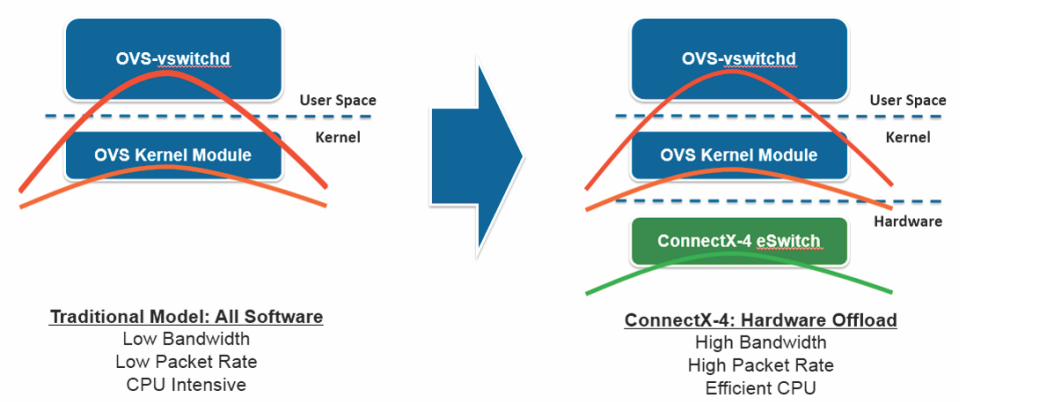


图 5. OVS 卸载可以释放 CPU 并提供性能优势

数据中心密度和异构性

为了提高效率，数据中心管理员不断探索能够提供更高的计算密度、更低的功耗和空间占用以及降低整体复杂度的创新设计。

Mellanox 多主机™技术在高数据中心密度和高效机架设计方面向前迈进了一步。多主机技术允许通过将适配器 PCIe 接口分隔为多个独立的 PCIe 接口，将多个计算或存储主机连接到单个互连适配器。

这样可以降低所需的适配器数量，减少线缆和交换机端口。

多主机技术还用于提供对架构的不同 CPU 或套接字直接访问，从而降低延迟和释放 CPU 周期。

多主机支持异构数据中心体系架构；连接到单个适配器的各个主机可以是 x86、Power、GPU、ARM、FPGA 或内存，从而可消除在计算元素之间传递数据或进行通信的任何限制。

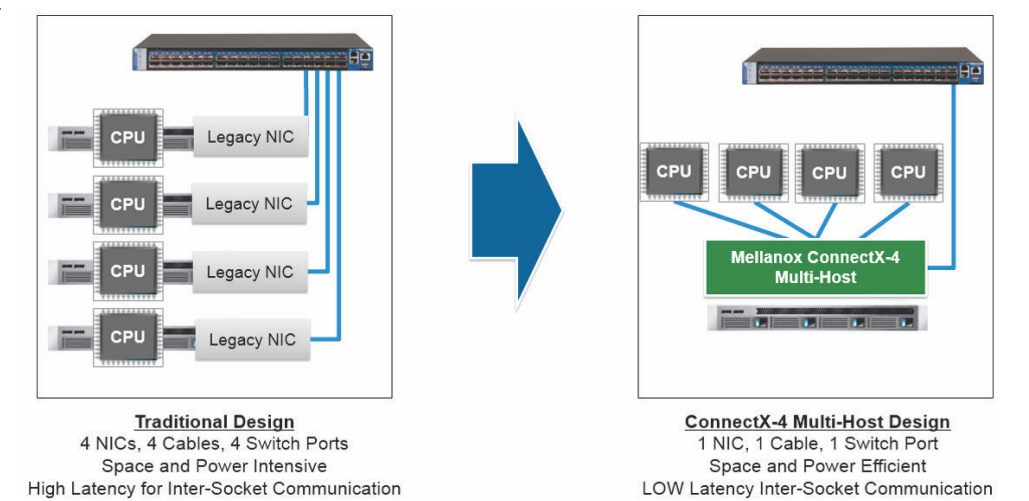


图 6. Mellanox 多主机节省了硬件和电源成本，并提高了性能

易部署

很多时候，云中的网络相关部署被忽视了，从而造成不必要的高部署和维护成本。为了使云操作能够顺利使用网络设备，应该考虑如下图所示的各个方面。

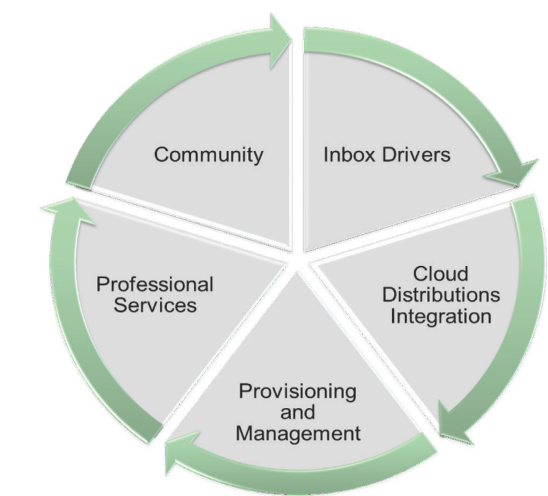


图 7. 易部署注意事项

Mellanox 云解决方案具备上游操作系统发行版和商用云产品的高级集成，再加上 Mellanox 的管理和自动化工具及专业知识。

这些优势使云管理员能够在 1 到 3 周的时间内使用标准的现成软件和硬件部署完整的生产环境。

内置驱动程序 — 在主要的发行版中内置了 Mellanox 驱动程序。如需其他必要功能，Mellanox OFED 驱动程序公开发布在 Mellanox 网站上。

http://www.mellanox.com/page/products_dyn?product_family=26&mtag=linux_sw_drivers

云发行版集成 — Mellanox 正在不断努力使其解决方案成为 OpenStack 发行版中的默认解决方案，如 Mirantis、RDO、Canonical 等。

预配置和管理平台 — 为 Mellanox 客户提供 NEO 管理平台，用于自动化云部署和预配置，并支持日常网络监控和操作。

NEO 具有丰富的 REST API，使其能够无缝集成到任何现有的管理框架中。

用于领先云平台的 CloudX 参考体系架构

专业服务 — 由于部署了大量不同类型的云（通过服务提供商、私有部署或 HPC），Mellanox 专业服务在帮助客户高效部署云环境方面拥有非常丰富的经验。

社区 — 对于那些想要自行处理部署和运营的公司，Mellanox 在 Mellanox 社区网站上提供了大量文档和社区讨论：

<https://community.mellanox.com/docs/DOC-2288>

为了最大限度地提高投资回报率并保护现有的高昂基础架构投资，网络解决方案参与整体云部署、预配置和运营流程与自动化至关重要。

OpenStack 集成

Mellanox 为各种 OpenStack 项目贡献了大量代码，用于：

1. 实现与网络设备相关的云管理流程的自动化
2. 优化其网络相关性能

目前提供以下集成：

- OpenStack Neutron 和 Nova — 用于云网络自动化
- OpenStack Cinder — 使用 iSER 的高延迟存储解决方案
- OpenStack Ceilometer — 用于网络计量

以下链接包含全面的参考体系架构文档：

<http://www.mellanox.com/openstack/pdf/mellanox-openstack-solution.pdf>

VMware vCloud 集成

使用 Mellanox 技术部署的 VMware 云可以利用 Mellanox ConnectX®-3 Pro（及更高版本）适配器的虚拟可扩展局域网 (VXLAN) 卸载功能，以提供 VXLAN 的可扩展性和效率，而不存在典型的相关不利后果（网络吞吐量大幅下降，CPU 利用率增加）。

此外，基于 VMware 的云可以从 Mellanox 高速适配器中受益，体现在实现虚拟机的高性能并增加 Hypervisor 上的虚拟机密度。

以下链接包含全面的参考体系架构文档：

http://www.mellanox.com/related-docs/whitepapers/CloudX_VMware_5_5_with_vShield_Reference_Guide.pdf

Windows Azure Pack 集成

Mellanox 使组织能够构建在 Windows Azure Pack 上运行的最高效的云。利用 SR-IOV、网络叠加 (Overlay) 卸载和 RDMA 等技术，使用 Mellanox 适配器构建的 Windows 云可实现最高水平的性能、效率和可扩展性。

Microsoft 的 Hyper-V Hypervisor 为租户流量使用 NVGRE 网络虚拟化。Mellanox 的 ConnectX®-3 Pro（及更高版本）网络适配器为其提供了有效的加速，可以减少由于网络通信造成的 CPU 开销，并产生更高的虚拟机密度。

在 Windows Server 2012 R2 中，Hyper-V 使用 RDMA/RoCE 进行实时迁移。该技术显著减少了虚拟机迁移所需的时间，特别是由于实时迁移造成的 CPU 开销。随着迁移成本大幅降低，运营商可以更快地平衡云流量负载。这使得集群能够以更少的资源储备运行，而不会损害任何 SLA。

Mellanox 适配器满足您的云需求

云存储使用由 Microsoft 存储空间上的 Microsoft 服务器消息块 (SMB) 协议 3.0 版增强的横向扩展文件服务器。SMB 3.0 文件服务器使用基于 RoCE 的 SMB Direct 来减少存储访问造成的 CPU 开销，从而确保最佳性能。

以下链接包含全面的参考体系架构文档：

[http://www.mellanox.com/related-docs/applications/Windows_Azure_Pack_\(WAP\)_CloudX_Reference_Guide_v1.0.pdf](http://www.mellanox.com/related-docs/applications/Windows_Azure_Pack_(WAP)_CloudX_Reference_Guide_v1.0.pdf)

Mellanox 适配器为云部署提供了最独特的优势组合，适用于广泛的需求和用例。无论优先考虑最高的原始数据速率、运行 HPC 应用程序、低延迟存储还是虚拟机密度，Mellanox 适配器都能提供当今市场上最先进和最高效的解决方案。

Mellanox 适配器针对云用例的优势	
10G/25G/40G/50G/100G 最高性能	✓
提高虚拟机密度	✓
在虚拟机中运行低延迟应用程序，如 HPC	✓
无需占用 CPU 即可加快存储速度	✓
利用智能卸载释放 CPU	✓
叠加 (Overlay) 网络无状态和封装/解封卸载	✓
提高机架密度并降低线缆和端口的复杂性	✓
与领先的云供应商（开源和商业）集成	✓
易于部署	✓
在大型云方面获得验证的现场解决方案	✓



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