

高效能OpenStack私有云解决方案



私有云发展的机遇和挑战

随着数据中心的不断扩容和业务系统的持续增加，越来越多的企业开始部署云计算平台，利用云计算的技术优势构建企业私有云，来解决传统数据中心面临的挑战和压力，包括成本压力、业务压力、运维压力。当前私有云已经步入了实质性的发展阶段，未来的3-5年将是云计算的快速发展期，数据中心建设将迎来一场变革。

然而私有云的部署和应用并非一蹴而就、简单易行，高效能的云计算系统应该可以为企业带来诸多收益，包括：提高资源利用率，降低总体拥有成本TCO；提升IT效率；资源集中、高效管控；提高业务稳定性；此外还要满足灵活部署、弹性扩展、按需服务等要求。私有云平台采用的架构和技术成为实现这些目标的关键。

高效能私有云

针对上述需求，Mellanox（迈络思）与UnitedStack联合发布了高效能OpenStack私有云解决方案，该方案的功能架构图如下：

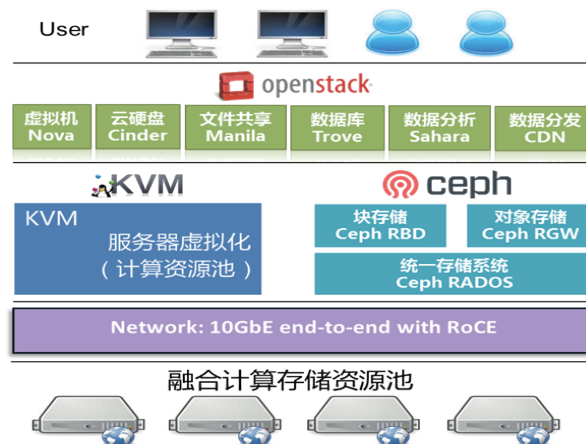


图1：UnitedStack私有云超融合架构

UnitedStack 私有云功能特性包括：

- 通过SDN 实现VDC 虚拟数据中心，保证业务的隔离和独立性
- 基于Ceph 的分布式高性能存储，极大地提升云主机的IO 性能
- 秒级的云主机性能监控，支持报警功能
- 基于安全组的云主机访问控制，灵活的安全组策略
- 强大的路由器，实现网关、端口转发、VPN 服务

该私有云完全采用OpenStack 开源平台和开发标准构建，100%完整支持原生OpenStack API。同时，UnitedStack 技术团队积极参与社区开发，所有源代码都会提交到OpenStack开源社区，保证所有代码开源可靠。

通过部署UnitedStack 私有云，用户的IT 环境将具有前瞻性，可随时适应不断变化的业务需求。用户可以保留灵活的选择，无需依赖特定的硬件、操作系统、应用程序或供应商，因此可以对在私有云中部署现在和未来的业务系统充满信心。

弹性网络互连

UnitedStack 私有云集群包含控制节点、网络节点、计算节点等，基于10G以太网构建。针对其网络需求和功能特点，Mellanox设计出基于VMS（Virtual Modular Switch）的聚合网络架构，配置灵活，动态扩展，并具备成本优势。

该网络架构包含三个配置选项，采用不同的交换机型号，用户可以根据规模的大小灵活选择。分别是小规模（小于160个节点），中等规模（160-240个节点），大规模（240个节点以上）。其中大规模配置的网络拓扑图如下：

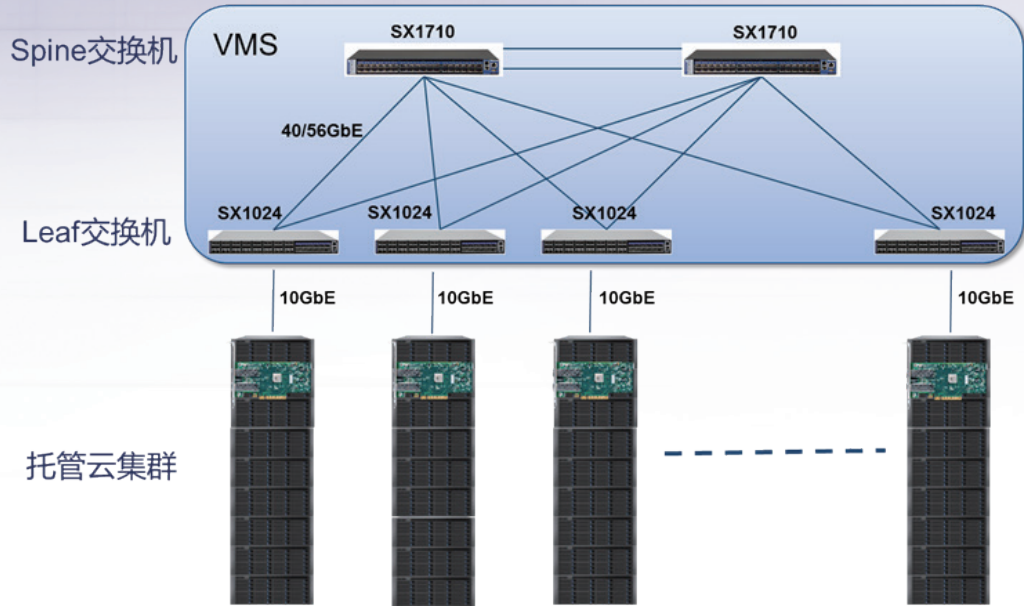


图2：基于Mellanox VMS的私有云网络架构

- Leaf交换机采用Mellanox SX1024，支持下行48个10G端口，上行12个40G端口（无阻塞）或者4个40G端口
- Spine交换机实现聚合网络通信，采用Mellanox SX1710，支持40/56GbE的网络连接，提供高达4.03Tb/s的无阻塞带宽
- 2台Spine交换机通过MLAG实现高可用
- Spine交换机和Leaf交换机支持40G互连，可以升级到56G
- 集群计算节点、网络节点、控制节点采用Mellanox ConnectX-3 Pro双端口万兆网卡，支持虚拟化和云计算加速

虚拟化加速

Mellanox的网卡支持广泛的虚拟化和云计算特性，包括通过网卡内置的嵌入式交换机（eSwitch）支持高性能SR-IOV，基于RDMA（远程内存直接访问）加速存储和虚拟机性能，以及针对Overlay Network的硬件加速和卸载。

Overlay Network（VXLAN/NVGRE/GENEVE）通过虚拟网络和物理网络的嵌套极大地提高虚拟网络的可扩展性，移动性，隔离性。然而基于Hypervisor的Overlay Network会显著加重系统的负荷和成本，因为数据包的处理（包括封装、解封装、分片、聚合、校验等）都需要消耗大量的CPU资源，降低系统的处理能力，最终造成网络吞吐量受限。

针对Overlay Network的性能瓶颈，Mellanox提供ConnectX-3 Pro和ConnectX-4系列网卡，通过硬件卸载引擎，将数据包的处理操作卸载到网卡上，从而极大的释放CPU资源，并提高网络吞吐量。

UnitedStack 私有云方案中已经基于Mellanox ConnectX-3 Pro 10GbE网卡测试和集成了VxLAN Offload功能，对比VxLAN no Offload的性能优势如下图：

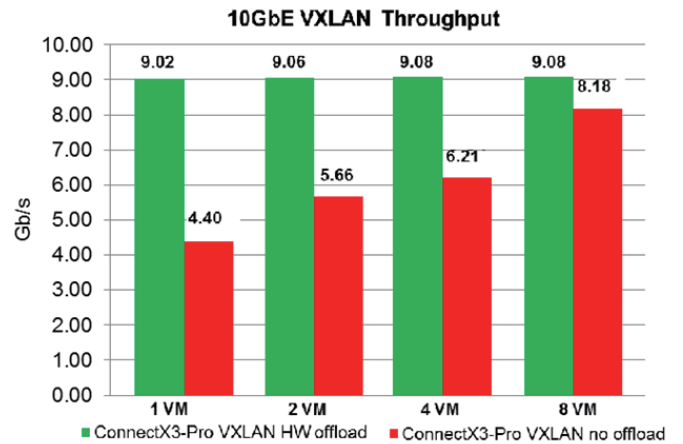


图3：VxLAN Offload显著提高网络带宽

Ceph存储加速

UnitedStack 私有云方案采用Ceph集群用于OpenStack Cinder块存储。云计算环境中的存储集群需要具备高可用性和高可扩展性，因此要求网络互连能够提供足够的带宽和灵活性来满足应用的需求。

Ceph的Cluster Network用于OSD互连，Public Network用于用户接入，针对数据类型和负载的差异，UnitedStack 私有云采用了不同的网络带宽。Public Network基于Mellanox 10GbE端到端网络

互连，而Cluster Network因为数据量较大，采用Mellanox 40GbE，主要承载OSD心跳，数据复制，负载均衡，数据恢复等。

该网络设计为Ceph存储集群带来了显著的性能提升，其带宽和IOPS对比1G和10G网络的优势如下：

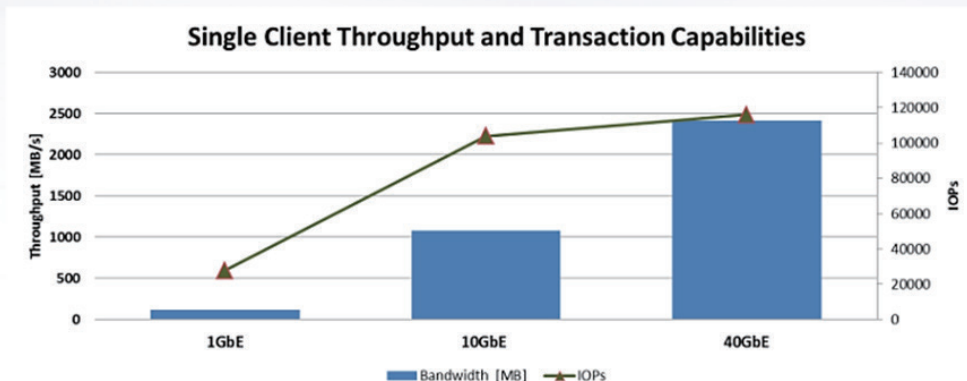


图4：Ceph集群基于40GbE网络的性能优势

方案部署和效益

基于Mellanox端到端以太网互连和虚拟化加速的UnitedStack 私有云方案在国内的私有云市场取得了巨大的成功，并大量部署，帮助用户提升了云平台的效率，得到用户的广泛认可。

该私有云方案的优势总结：

- 开源Openstack：基于OpenStack开源平台和标准API，支持新版本及时更新和在线升级；
- 弹性网络：网络拓扑基于Mellanox VMS架构，配置灵活，动态扩展，并极大地降低硬件成本；
- 高性能：Mellanox以太网核心交换提供40/56Gbps的超高带宽，实现Tb级别的聚合能力，以及纳秒级别的超低延迟；
- SR-IOV：Mellanox网卡内置嵌入式交换机eSwitch，实现本地虚机通讯，服务质量QoS，拥塞控制，隔离和安全；
- Overlay Network Offload：VXLAN/NVGRE/GENEVE硬件卸载，降低CPU负荷，提高网络吞吐量；
- RDMA加速：Mellanox网卡支持RDMA硬件加速，提升应用性能的同时，显著提高CPU资源的利用率；
- Ceph集群：基于40GbE网络的Ceph集群，为云主机提供更高的带宽和IOPS；
- 线性扩展：支持集群在线线性扩展，满足业务升级和扩容的需求。



北京迈络思科技有限公司

咨询电话：+86-10-57892000

销售咨询：china_sales@mellanox.com

市场合作：marketing_cn@mellanox.com

*欲了解更多欢迎登陆www.mellanox.com

