云计算实战：浅谈私有云环境该如何进行网络规划

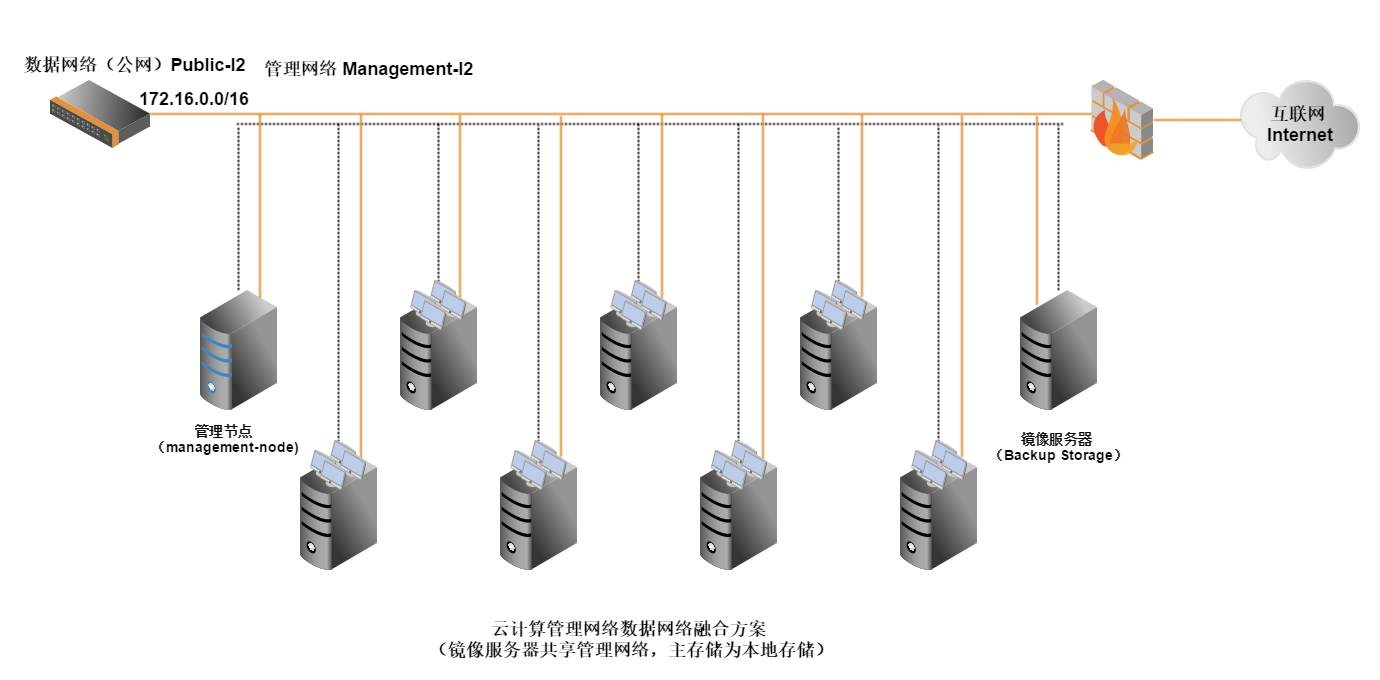
by Yongkang You on Feb 4, 2016

前言

随着私有云管理软件（如ZStack）的普及，越来越多的用户可以利用这些软件快速地搭建私有云平台。 不过在搭建私有云之前，大家都会想：该如何在私有云环境中进行网络规划？ 是不是把所有的物理机都连接到交换机，给它们分配一个IP地址就好了呢?

目前，用户可以利用现有的网络环境快速地搭建一套私有云环境，然而在考虑私有云的隔离性、稳定性、可用性等需求时，应对网络有更好的规划。 今天我们就来谈谈，私有云环境里如何进行网络规划，希望能为计划构建私有云的企业提供一套有效的方案。

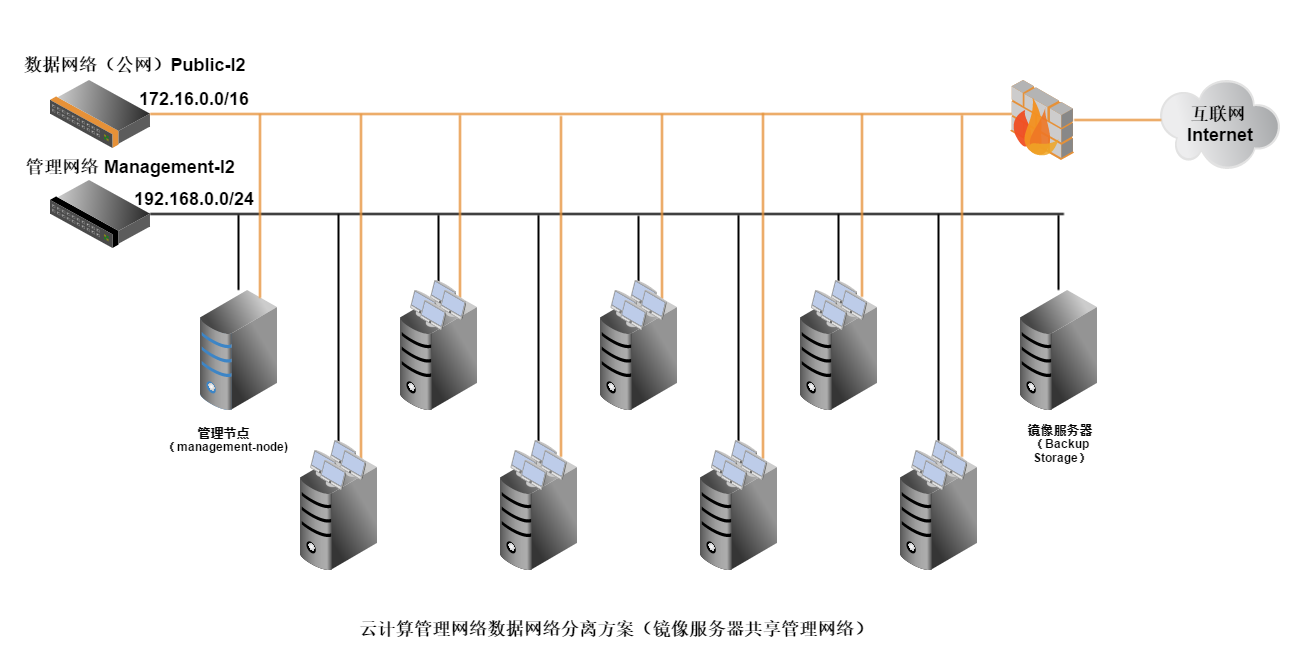
基本要求：管理网络和数据网络分离



通常很多企业的服务器都配有超过一块网卡，但是只连接和使用了其中一块。在搭建私有云的过程中， 很容易把网络的数据流量和私有云的管理流量全部使用相同的网卡上。 虽然这样的配置可以工作，但是会存在两个显著的问题：

1. ZStack会把需要控制安全访问的连接限制在管理网络，例如Virtual Router的网络控制，云主机的VNC连接。 合并管理网络和数据网络后，就会把限制在管理网络上的安全访问暴露到公网上。
2. 云主机的网络流量占满网卡的带宽时，ZStack 管理节点无法控制和操作云资源。

如果只解决第一个问题，我们可以采用给网络划分Vlan的方式，把管理网络和数据网络划分到相同网卡不同的Vlan上。 但是Vlan的方式无法解决第二个问题。所以我们建议用户可以在演示环境中使用单网卡模型快速熟悉私有云的部署。 但是在生产环境中，至少需要把管理网络和数据网络分离。



在ZStack实际部署过程中，如何将管理网络和数据网络分离呢？ 首先，我们需要保证管理节点和物理机上都至少配置了两块网卡，连接到了不同的网络，并且设置了对应的IP 地址。

其次如果用户使用的是带虚拟路由器（Virtual Router）的网络模型，例如EIP、 [带虚拟路由的扁平网络](http://zstack.org/tutorials/flat-network-ui.html)， 那么就在创建二层网络（L2 Network）的时候，除了添加一个名为Public-L2、 IP地址空间为172.16.0.0/16的网络外，还需要添加另外一个名为Management-L2， IP地址范围为192.168.0.0/24的管理网络 （需要将服务器已经占用的IP地址剔除。例如管理节点，物理服务器的IP地址是从192.168.0.1～192.168.0.20， 那么设置的可以用IP地址范围可以为192.168.0.21～192.168.0.254）。 在设置虚拟路由器规格（Virtual Router Offering）的时候，分别把Public-L3 和Management-L3设置到正确的L3位置。 另外在添加物理主机的时候，物理机主机的管理IP需要指定使用物理主机上在192.168.0.0/24网段中对应的IP地址。

如果用户使用的是[不带虚拟路由的扁平网络](http://zstack.org/tutorials/flat-network-ui-no-vr.html)。 那么无需特别添加名称为Management-L2的管理网络，只需要添加一个Public-L2的网络。 指定ZStack管理节点通过管理网络控制物理服务器的方法是，在添加物理服务器的时候， 特别指定使用物理主机上在192.168.0.0/24网段中的那个IP地址。

另外，如果数据网络和管理网络均为千兆网络的时候， 我们还可以让镜像服务器和物理服务器之间的云主机镜像传输的通道与管理网络共享，以节省数据网络的网络流量。 用户只需要在添加镜像服务器（备份存储服务器）的时候，使用镜像服务器上网段在192.168.0.0/24上对应的IP地址。

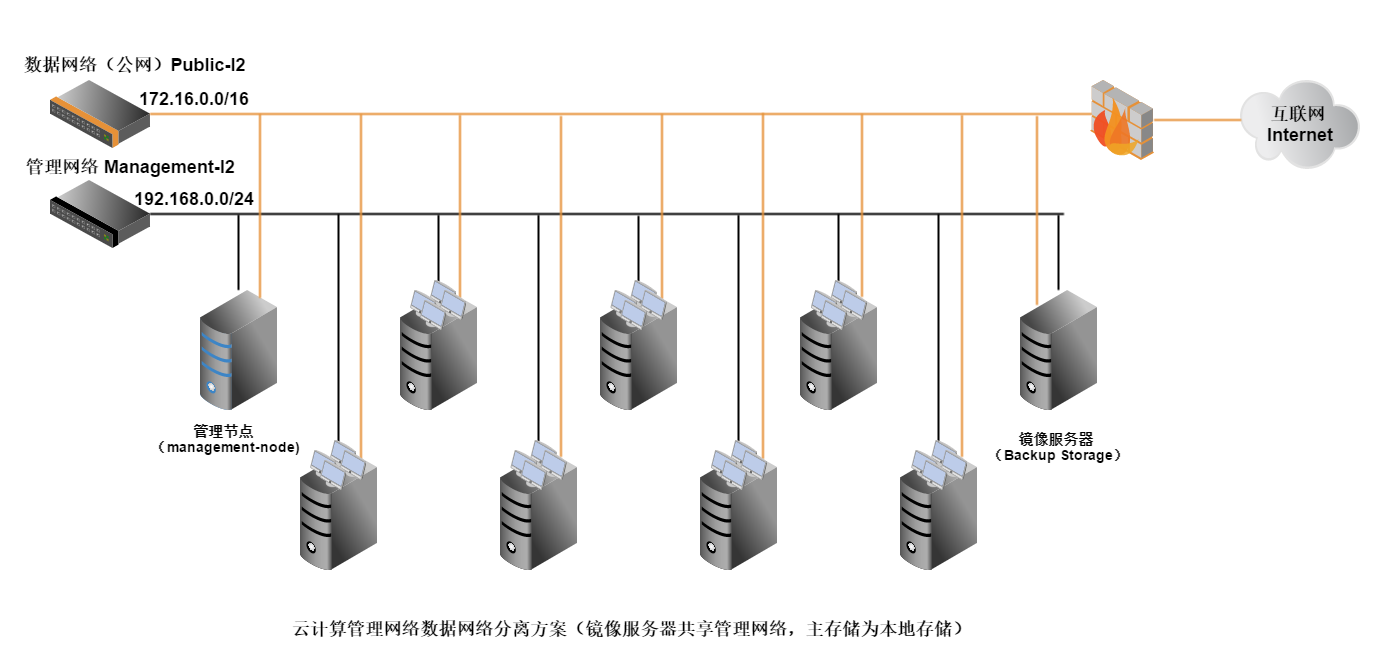
在生产环境中，为了保证交换机以及网络的高可用性，往往还会对数据网络构建bond模式。这个时候可能会使用双网卡。

下面我们来介绍几种典型的网络规划场景：

场景一：企业内部小型数据中心（如100台服务器）

市场调查表明，目前国内大部分企业的数据中心规模都在100台服务器以下。网络以千兆网络为主，少量万兆网络， 少量依然是百兆网络。存储多以RAID模式的本地存储为主，辅以商业存储（网络共享存储）作为数据备份。 对于这类企业来说，将现有的数据中心迁移到私有云的架构，其实是非常容易的。在充分的利用现有资源的情况下， 可以快速地提高企业的生产效率。

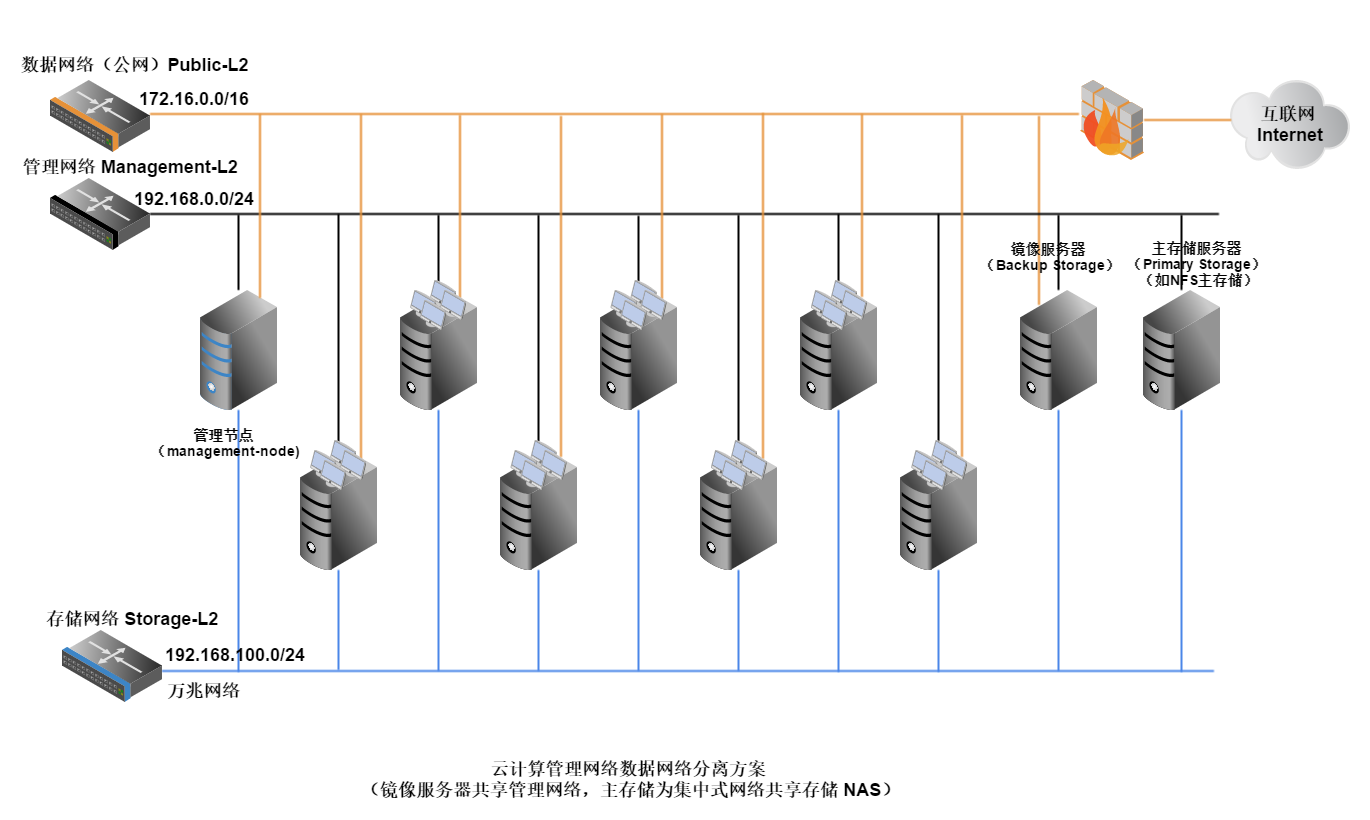
由于主要采用本地存储和千兆网络环境。我们推荐使用：私有云中的扁平网络模型+本地存储。 这种模式的好处是，最大化地利用现有资源。既提升性能和安全性，又能快速实施。



在上一幅图中，我们可以看到两个网络，分别连接了ZStack管理节点，物理服务器，镜像服务器。 ZStack管理节点通过192.168.0.0/24的网络连接并控制物理服务器和镜像服务器。 物理服务器上的云主机则通过172.16.0.0/16的网络进行数据交换，透过企业的防火墙访问Internet。

场景二：使用NAS的企业数据中心

与场景一的环境不同，企业网络可能已经升级到万兆网络，并且购买了提供了NFS的网络访问的商业存储。



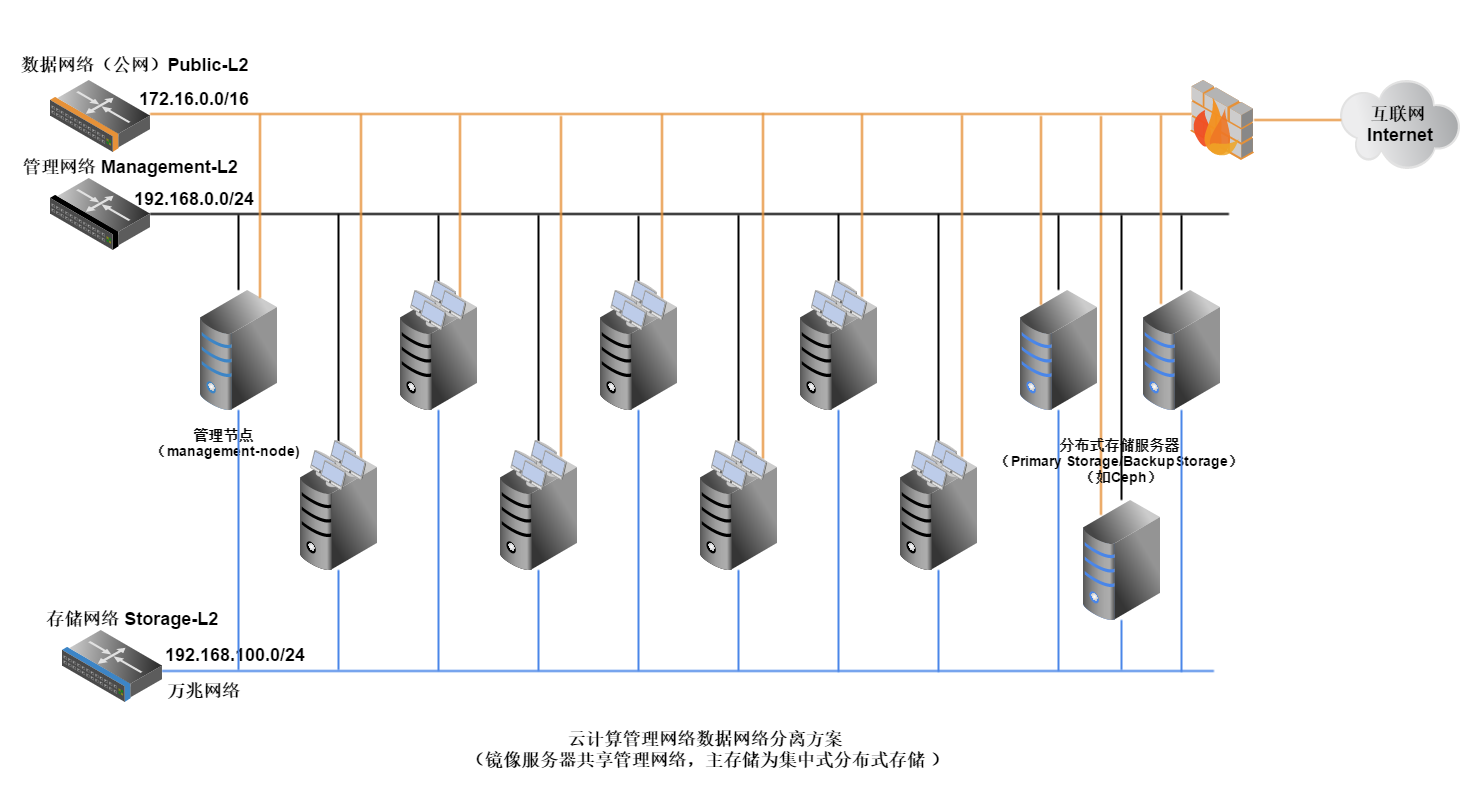
这种场景过渡到私有云也非常的容易！只需在选择主存储的时候，选择NFS的方式。 另外由于使用了网络共享存储，云主机的磁盘IO访问都会通过网络访问NFS主存储。 因此单一的数据网络很可能无法满足大量的磁盘读写请求。我们需要新划分一个存储网络（万兆）。 为什么需要万兆呢？因为云主机的云盘都是放在NFS存储中，如果只是千兆的网络环境， 即使仅有一个云主机在进行磁盘操作，最多也只能使用到1gbps的带宽，理论磁盘读写的上限为125MB/s。 考虑到多云主机以及网络传输效率，千兆网络将会极大的限制云主机的磁盘IO性能。

在使用存储网络的时候，我们需要先将物理服务器，NFS主存储，镜像服务器，管理节点上空闲的网卡连接到新的存储网络， 并且配置存储网络的IP地址（例如192.168.100.0/24网段）。在添加NFS主存储和镜像服务器的时候，使用这些IP地址即可。

场景二的优势是，计算和存储分离！云主机可以在线迁移，一旦计算节点失效，云主机可以快速恢复。 商业NFS主存储在数据的高可靠、高可用、性能方面都有较好的表现。 当然采购商业NFS主存储的价格会是一笔不小的费用！再考虑到HA模式需要对万兆存储网络进行双网络的配置， 整个网络的投资也将提高不少。

场景三：使用分布式主存储的企业数据中心

由于场景二中描述的商业NFS主存储可能会有额外的较大开销，以及考虑到目前流行的软件定义存储和分布式存储， 一些有较强运维能力的公司开始尝试使用Ceph来搭建存储中心。



这种模式也需要用户拥有万兆的网络环境，而在ZStack中搭建分布式存储网络模型，则和NAS架构相比， 仅需在添加主存储时，选择使用Ceph，并且指定192.168.100.0/24网段Ceph的MonURLs。

使用Ceph的好处有：成本相对低廉（完全使用免费开源软件搭建，且不考虑聘请额外的Ceph运维人员）， 数据“没有”单点故障（要求运维良好），数据相对安全（多副本存储于不同的机器）。 注意到当前市场上以有一部分公司都已经在生产环境中使用了Ceph。 不过考虑到目前的市场反馈，保证Ceph稳定性，如果中小公司打算采用Ceph方案， 建议请联系提供专业Ceph解决方案的厂商，如XSKY。

另外，细心的读者可能会问，我们为什么图中使用计算和存储分离的Ceph架构，而不是计算和存储融合的架构。 这个地方主要也是考虑到Ceph在大IO（例如计算节点上云主机较多，且IO操作比较频繁）的时候，会消耗大量的物理机CPU。 如果采用融合的架构，如果不进行有效的隔离，可能会导致云主机的效率降低。当然，具体采用何种分布式架构， 还需要客户根据自己的场景来进行规划，也可以咨询专业的Ceph厂商来提供建议。

总结

至此，我们介绍了三种网络架构，分别适用于本地存储、NAS存储、以及分布式存储。 希望这篇文章可以帮助使用ZStack的用户快速的构建稳定可靠的私有云环境！

对ZStack和相关私有云部署感兴趣的用户，还可以加入ZStack QQ群（410185063）获取更多的帮助。