联通虚拟资源云存储平台设计与实现.ashx

联通虚拟资源云存储平台的总体架构，由四个部分组成，分别是虚拟资源管理、系统管理、基础设施管理以及虚拟化控制层。如图5-1所示，是联通虚拟资源云存储平台的总体架构图。图5-1 总体架构图



1）虚拟资源管理虚拟化资源管理，是实现整个平台资源管理调用的关键层。对外提供统一接口，可同时调用管理多个不同类型的虚拟资源池。向上屏蔽资源使用复杂性，向下屏蔽异构虚拟化系统调用，使用户操作虚拟资源如物理资源般方便。该层以域、资源组、资源池、集群等概念实现物理资源的逻辑分配，保证资源的弹性化使用及资源安全可靠。2）系统管理系统管理，是实现平台的整体业务管理。由用户认证、主控台管理、用户管理、租户管理、日志消息管理五部分组成。用户认证登陆系统之后，在主控台展示当前用户所在租户的资源信息，了解目前的数据情况和使用情况。通过租户对万方数据

第五章系统开发与实现37于资源进行有效的划分和管控，提供日志消息管理服务，可以清楚直接的追查资源的使用历史和使用情况。3）基础设施管理基础设施管理实现联通大规模资源分区域管理并向上提供统一控制API。通过分层收集不同区域中物理节点、存储、计算和网络等设备的基本信息和状态信息，实现分区域的资源获取及资源调度，向上提供灵活高效的资源管理接口。4）虚拟化控制层虚拟化层是联通虚拟资源云存储平台中承上启下的模块，完成物理资源向虚拟资源的转化。通过从基础设施层获取物理资源基本信息，利用OPENSTACK模块，实现相应物理资源的虚拟化，其中包括网络资源虚拟化、计算资源虚拟化和存储资源虚拟化；并将不同的虚拟化资源组成基础资源服务，向上提供池化服务

5.1虚拟资源管理虚拟资源管理（简称VMS）该模块设计用于满足用户需求集群定义、负载均衡、弹性伸缩。VMS系统主要实现虚拟资源的分配及管理。通过域、资源组、资源池、集群的业务单元划分，实现面向多用户的可管理、可配置、可调度的资源需求。该平台向上提供虚拟资源服务接口，向下屏蔽各种虚拟化技术，此次项目使用OPENSTACK虚拟化平台，后期支持扩展其它虚拟化平台。

5.2 系统管理系统管理主要是实现联通的用户管理和消息管理两部分，用户管理主要用于管理平台的用户及其状态和基本信息；消息管理主要用于系统的业务消息和功能消息部分的管理。已维持系统的基础管理部分。

5.2.1用户认证

5.2.2主控台管理

主控台是对于用户最关心的数据，进行数据统计。主要统计数据由：虚拟机物理机、虚拟内核、数据卷、路由器、安全组、浮动IP、内存容量、数据卷和数据卷快照总容量八部分组成。

5.2.3用户管理

5.2.4租户管理

租户管理，主要是对整个平台中的租户进行基础数据管理的服务。在本论文2.1.7章节中，介绍过OPENSTACK的身份认证服务Keystone，其中就提到过租户（Tenant）这个概念。租户这个概念，本身是来源于OPENSTACK对于资源的管理所划分出的业务的管理单位。也就是最小的管理群，资源通常以一个租户为单位进行操作和共享。租户管理，以表格形式展示，由租户名称和操作两部分组成。是平台管理员对于平台租户进行可视化管理的界面，租户管理界面如图5-10所示。就租户自身而言，就是一个组群的概念，将平台用户管理中的用户，进行以资源分配为目的的划分。租户和用户之间是多对多（N:N）的关系，也是一个租户可以有多个用户，一个用户也可以有多个租户。而用户也只能操作和共享所在的租户的资源。

租户管理，提供刷新数据、设置配额、管理租户和快速检索租户四个功能点。刷新租户数据，是为了方便存在多个平台管理员或者于多处同时登录时，能快速同步租户信息。设置配额，是指对于当前租户的资源分配设置一定的配额。界面如图5-11所示。设置配合，在实现上使用的是模态框模式，是对当前租户下的资源进行配额设置的功能。对虚拟机和物理机数量、虚拟机内核数量、内存容量、数据卷个数、数据卷总容量、路由器个数、浮动IP个这数、数据卷快照个数、端口个数、子网个数，这十种资源的相关参数进行配置。租户管理，主要是为当前租户设置所属的用户，也可以理解为把用户选择进不同的租户，一个用户若是不属于任何一个租户，那么他是没有资源可进行操作的。而快速检索租户，是为了方便平台管理员在租户管理界面，可以快速找到自己想要操作的租户信息。

5.3基础设施管理该模块设计是针对于联通集团公司内部的物理资源，进行统一管理和控制的功能。管理的物理资源主要包括三部分，计算资源、存储资源、网络资源。联通虚拟资源云存储平台的基础设施管理主要实现区域级别计算、存储和网络等设备的分层管理。通过层级资源获取及级联资源调度，实现大规模资源分区域管理并向上提供统一控制API。保证云平台具备注册新的云区域、管理云区域（删改查）、查看所有区域的概要情况、以及根据区域进行网络资源和存储资源管理的能力。基础设施管理子系统逻辑分为三层：区域、域、资源组。资源组是具备完备云能力的一组物理计算机；域实现不同组间的资源管控，且只控制组间调度；通过层层细致的分层管理框架，能够根据具体的业务需求定制灵活高效的云区域资源管控机制，实现大规模的云资源管理。

电子科技大学硕士学位论文50图5-12 系统管理-日志消息管理界面5.3基础设施管理该模块设计是针对于联通集团公司内部的物理资源，进行统一管理和控制的功能。管理的物理资源主要包括三部分，计算资源、存储资源、网络资源。联通虚拟资源云存储平台的基础设施管理主要实现区域级别计算、存储和网络等设备的分层管理。通过层级资源获取及级联资源调度，实现大规模资源分区域管理并向上提供统一控制API。保证云平台具备注册新的云区域、管理云区域（删改查）、查看所有区域的概要情况、以及根据区域进行网络资源和存储资源管理的能力。基础设施管理子系统逻辑分为三层：区域、域、资源组。资源组是具备完备云能力的一组物理计算机；域实现不同组间的资源管控，且只控制组间调度；通过层层细致的分层管理框架，能够根据具体的业务需求定制灵活高效的云区域资源管控机制，实现大规模的云资源管理。5.3.1 虚拟机管理虚拟机资源管理，主要是用户对于自己租户下所属的虚拟机进行管理控制。通过管理员创建虚拟机，在虚拟机服务内定义主要参数到达资源规划、调度，简化用户穿件流程的目的。界面如图5-13所示。

虚拟机服务，是以平铺模式展现，用户在进入界面后，能清晰看到当前自己能操作管理的虚拟机及其情况。虚拟机服务管理提供了查看虚拟机信息、查看虚拟机对应实例信息、修改基本数据、克隆虚拟机、开关机启停服务和删除虚拟机，这六个功能服务。查看虚拟机信息，主要是对虚拟机数量、状态、类型进行平铺展示。查看虚拟机实例，“实例”这个概念在本论文的2.1.1中，介绍OPENSTACK的Nova服务时详细介绍过。“实例”就是一个服务资源单位，对应到本章节，就是一台虚拟机。但是是一台包含业务属性的虚拟机。因而，查看虚拟机对应实例信息功能，是将这台虚拟机与之对应的实例信息进行数据读取，并展示。修改基本数据，是对虚拟机名称进行修改的功能，虚拟机ID为唯一标识符，而名称主要是业务属性，

因此可以依据用户的业务需求进行修改。克隆虚拟机，主要是通过单台虚拟机创建多台虚拟机的功能，OPENSTACK的Nova服务，会先对需要克隆的虚拟机进行备份，然后再对备份的虚拟机进行克隆。开关机启停服务，是指对虚拟机进行开关机操作，只有开机状态的虚拟机，才可以被用户所看到和操作，克隆虚拟机只有在关机状态下，才能进行这一操作。删除虚拟机，就是对虚拟机进行删除操作，从OPENSTACK底层库里对Nova的虚拟机表进行这条数据的删除处理，以及相关实例和信息的清除。

5.3.2 数据卷管理数据卷资源管理，主要是用户对于自己租户下所属的数据卷进行管理控制。通过管理员创建数据卷，在数据卷服务内定义主要参数达到资源规划、调度，简化用户穿件流程的目的。

数据卷服务，是以平铺模式展现，用户在进入界面后，能清晰看到当前自己能操作管理的数据卷及其情况。数据卷服务管理提供了查看数据卷信息、查看数据卷机对应实例信息、修改数据卷、克隆数据卷、开关机启停服务和删除数据卷，这六个功能服务。查看数据卷信息，主要是对数据卷数量、状态、类型进行平铺展示。查看数据卷实例，“实例”就是一个服务资源单位，对应到本章节，就是一个数据卷。但是是包含业务属性的数据卷。因而，查看数据卷对应实例信息功能，是将当前数

据卷与之对应的实例信息进行数据读取，并展示。修改数据卷，是对数据卷名称进行修改的功能，数据卷ID为唯一标识符，而名称主要是业务属性，因此可以依据用户的业务需求进行修改。克隆数据卷，主要是通过单个数据卷创建多个数据卷的功能，克隆数据卷，也会先进行数据卷备份，再对其已经备份好的数据卷进行克隆操作。开关机启停服务，是指对数据卷进行开关机操作，只有开机状态的数据卷，才可以被用户所看到和操作。删除数据卷，就是对数据卷进行删除操作，从OPENSTACK底层库里对Volume的数据卷表进行这条数据的对应的删除处理，以及相关实例和信息的清除。5.3.3 资源池管理资源池管理，主要是用户对于自己租户下所属的资源池进行管理控制。每个资源池可以定义不同的调度策略，系统可以根据不同的调度策略选择不同的资源集群和虚拟化类型，界面如图5-15所示。资源池管理服务，采用平铺模式展现资源池。用户在进入资源池管理界面后，能清晰看到当前自己能操作管理的资源池信息及其情况。资源池管理提供了查看资源池信息、修改资源、启停服务和删除资源池，这四个功能服务。查看资源池信息，主要是对资源池基本信息进行平铺展示。包括：资源池名称、资源池状态、资源池类型、虚拟化类型、可用域、数据卷类型以及网络，七个部分。修改资源，是指对可以查看到的资源池信息，进行修改操作。启停服务，是指对资源池进行启停操作，只有可用状态下的资源池，才可以被用户所看到和操作。删除资源池，就是对资源池进行删除操作，从OPENSTACK底层库里进行这条数据的对应的删除处理，以及相关信息的清除。5.4虚拟化控制层该模块支持OPENSTACK，满足用户基于应用提出的基础资源能力需求，包括应用基本管理、应用迁移、应用发布、应用集群、应用容错。同时主机管理能力、虚拟主机容错性能也是该模块设计着重考虑的性能问题。虚拟化层VM-Layer主要实现基础设施资源的虚拟及池化管理，基于OPENSTACK实现。该模块通过将一堆物理设备以资源服务（计算资源池、网络资源池、存储资源池、控制资源池）的形式展现给用户，使用户从繁琐的基础资源维护工作中解放出来，专心关注业务应用系统。该系统的资源池化分别基于OPENSTACK组件实现：计算资源池基于Nova组件，实现虚拟机相关操作；网络资源池基于Quantum组件，实现虚拟网络配置及管理；存储资源池基于Cinder组件及GlusterFS实现，

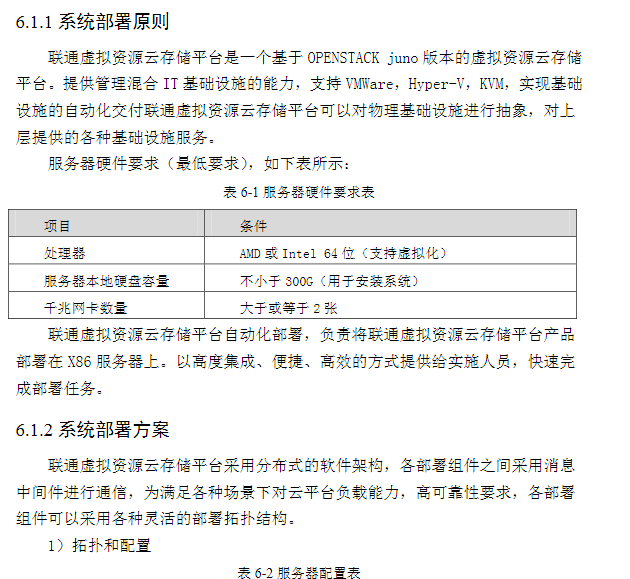
为整个联通虚拟资源云存储平台提供存储服务。

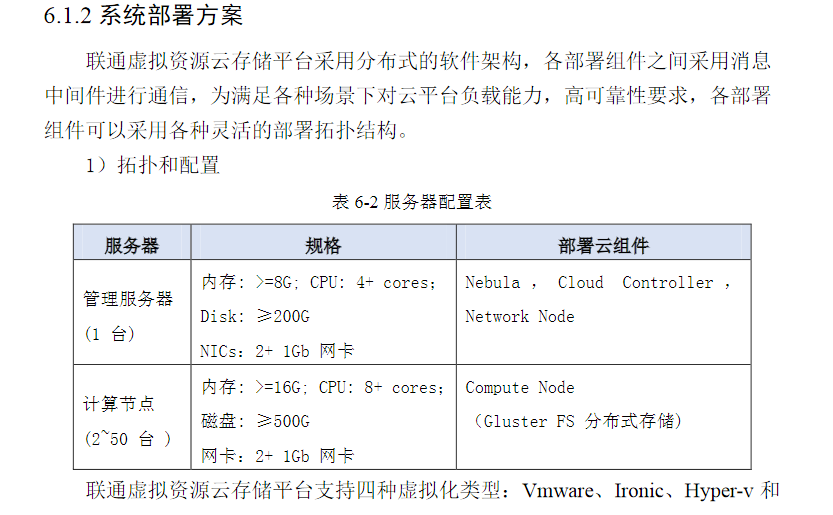
5.4.1 计算资源池本节功能包括虚拟机实例创建、虚拟机实例生命周期管理、虚拟机实例使用、虚拟机模板管理、虚拟机操作系统管理、虚拟环境管理、云主机。计算资源池是虚拟机的承载体，负责管理虚拟化层，主要任务是对外提供满足使用要求的虚拟机服务，可按用户要求提供单机、双机和多机集群应用能力。该模块基于Nova组件实现，通过可靠消息队列或直接调用与网络、存储等模块交互，实现租户需求资源平台构建。可配置实例相关策略，包括配置实例相关网络资源、配置实例相关存储资源的能力。同时也是上层API对于虚拟机操作的直接执行者，包括对虚拟机实例进行启动、停止、迁移、重启以及释放。

5.4.2 网络资源池该节设计用于满足用户需求IP管理、区域分段、功能划分、对外连接、网址映射。网络资源池是由可灵活分配的一堆虚拟网络组成。面向一个租户的资源申请，虚拟网络规划将租户对计算、存储资源的需求整合起来，以服务形式提供给租户，保证了资源的分配、回收及安全隔离。该系统基于Quantum组件实现，通过Quantum组件构建网络节点，实现计算节点、虚拟机、外网的基于策略互通访问。主要功能包括虚拟网络管理、IP地址分配、虚拟网络设备的管理，虚拟网络设备的访问控制以及联通虚拟资源云存储平台内的虚拟设备与外部网络之间的相互访问。

5.4.3 存储资源池该节设计用于满足用户需求存储空间管理、存储访问能力、存储卷实例、冗余备份、配额管理。同时存储系统访问性能、存储系统容错性能也是该模块设计万方数据

第五章系统开发与实现59着重考虑的性能问题。存储资源池为整个联通虚拟资源云存储平台提供存储服务。这些服务包括：为管理节点提供镜像与模板存储服务；为计算节点提供实例存储服务；为虚拟机系统卷提供查询服务；为虚拟机提供数据卷添加、删除、挂载、卸载、快照等服务；为虚拟机提供备份还原服务。





联通虚拟资源云存储平台支持四种虚拟化类型：Vmware、Ironic、Hyper-v和

Kvm，但“联通虚拟资源云存储平台压缩包”暂只支持Vmware、Ironic和Kvm三种虚拟化类型的自动化部署，Hyper-v需手动安装（详见本文6.1新增hyper-v计算节点）。现阶段，一台计算节点在同一时刻只能开启一种虚拟化服务。2）存储系统采用Gluster FS 分布式存储，可以在计算节点之外单独部署Gluster FS 集群，也可以在计算节点融合部署Gluster FS。根据应用对I/O性能的要求，可以加入SAN存储到存储资源池。联通虚拟资源云存储平台自动化部署存储方案暂不支持混合存储方案，并且支持的文件系统只有Ocfs2和Glusterfs两种。现阶段，需挂载存储的目录有三个：镜像目录(/var/lib/glance/images），存储目录（/var/lib/cinder/volumes）和虚机目录（/var/lib/nova/instances）。其中，cc节点会挂载这三个目录；nn节点不会挂载任何目录；nc节点会根据不同存储挂载不同目录，Ocfs2会挂载虚机目录和存储目录，Glusterfs只会挂载虚机目录。3）网络所有管理和计算节点连接到同一物理网络，使用VLAN 划分二层网络。可以将云平台管理网络流量/虚拟机应用流量/存储流量做网络隔离。



基于openstack的私有云管理平台的设计与实现

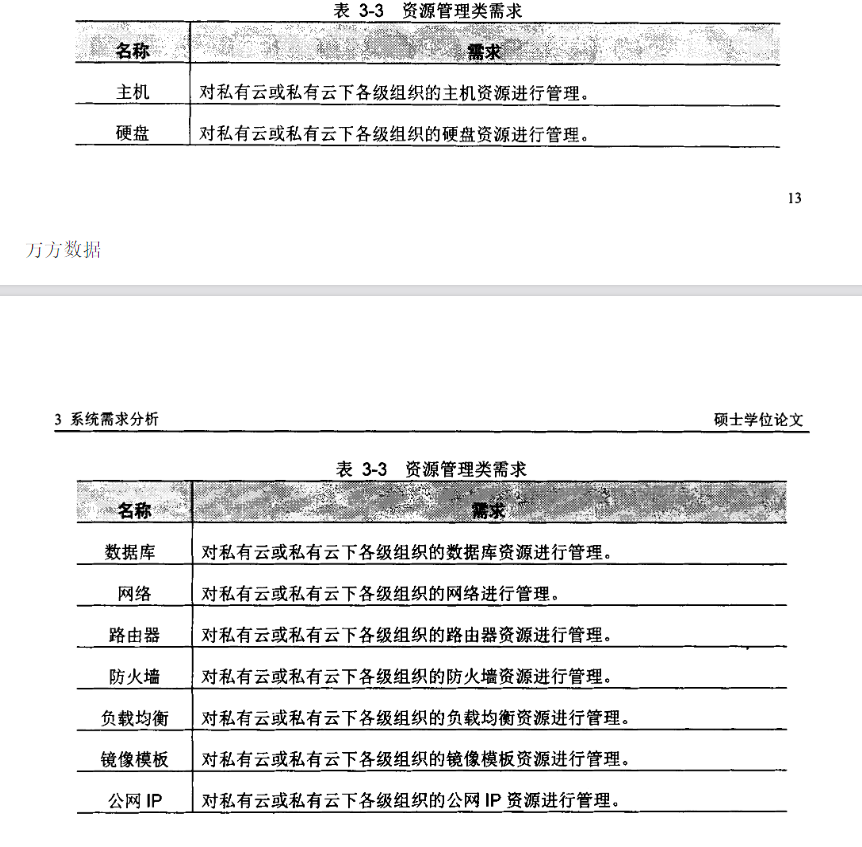
存储虚拟化的目标是整合异构的存储并将零散的存储资源整合起来，将存储的物理实体与逻辑表示分离开，以达到存储资源的动态调度的目的。使用者不会受到异构存储设备的困扰，只需利用分配给他们的逻辑卷，就可以得到大容量、高数据传输性能的存储系统。存储虚拟化有三个层面的实现方式，即主机或服务器层、存储网络层和存储设备层。在实际使用中，既可以单独使用，也可以在多个层面共同实现【“】。1)基于主机或服务器的虚拟存储方案基于主机或服务器的存储虚拟化的实现主要是通过主机的逻辑卷管理软件实现的，通过逻辑卷管理软件可以将分配给主机的逻辑卷进一步虚拟化，做到对多个逻辑卷的统一管理与配置。这种方式在主机系统和服务器上应用较为广泛，但是由于企业信息规模较大，对系统的性能和稳定性等有较高的要求，故而较大规模的私有云管理平台不适用这种虚拟化技术。2)基于存储设备的虚拟存储【删方案当多个服务器访问同一个磁盘阵列时，基于存储设备的存储虚拟化技术将是一个好的选择。基于存储设备的虚拟化存储主要是在存储设备的控制器上完成的【45】，它对其所管理的存储通过在控制器上添加虚拟化功能，将存储设备的存储容量划分为多个存储空间(LUN)，来提供给不同的主机系统。对于较小规模的企业，存储数据量较小、设备较简单的情况下，适用于该种方式实现的虚拟化实现方案。3)基于存储网络的虚拟存储方案这种虚拟化技术主要是在存储局域网的内部完成的，通过在局域网中添加虚拟化引擎实现。这是近年存储行业比较热门的一个发展方向。以上的两种存储的虚拟化方法优点都可以在其上同时体现，该方式可以支持数据中心级别的存储管理以及异构主机系统的存储系统。基于存储网络的虚拟化存储主要是在存储域网络(SAN)内完成的，可以通过在网络交换机上安装虚拟化软件实现，也可以是特有的虚拟化设备。它不仅可以完成对异构的主机系统和存储系统的管理，还可以支持数据中心级别的存储管理。对于数据量大的企业或省部级的政府部门，往往采用多种类型的主机和存储设备故而适合采用这种基于存储网络的虚拟化技术。

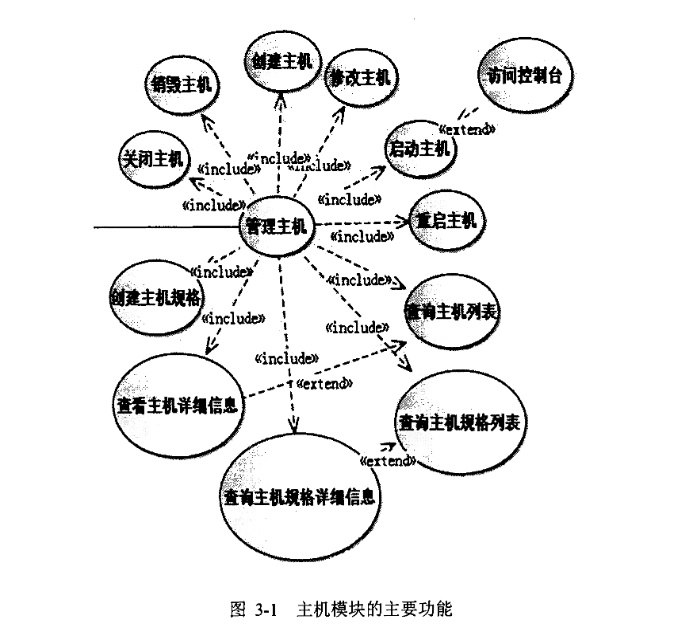
虚拟化技术具有资源共享、负载动态优化、自动化管理、安全性、节省资金绿色环保、解决平台依赖问题的优势。1)资源共享我们可以利用虚拟化技术，将包括存储、服务器、网络等资源整合起来，从而可以提高资源的利用率，减少资源的浪费。2)负载动态化动态优化主要有两个方面的内容：一方面资源的供给可以随着业务系统的工作负载动态的调整，虚拟化的出现使得这一功能得以实现；另一方面，通过使用虚拟化技术使得资源可以动态的调整，做到及时的部署与释放，提高整个企业IT系统资源的利

用率，减少物理机器的数量。3)自动化管理我们可以通过虚拟化技术将丰富多样的基础设施、中间件、操作系统等整合成一个虚拟的资源池，从而在使用时可以屏蔽底层硬件的差异，做到统一管理[49】。此外，还可以将企业常用的服务、软件等做成一个虚拟化的组件模板，从而可以实现快速的部署和自动化的管理，提高管理工作的效率。4)安全性因为虚拟化技术，我们可以在任何时刻极其快速的对运行的系统进行快照的制作与管理，快照的存在使得服务器收到恶意的攻击时也可以很方便的恢复，很好的保证了系统在突发情况下的安全性。5)绿色环保随着企业IT系统规模逐渐庞大，在软件、硬件、散热等方面的投入也随之增长。为了保证系统的稳定运行，在业务增加时，服务器数量及类型选择等有了一定的局限性。虚拟化技术的出现，使得底层资源的利用率大大提高，降低了整个基础架构的总成本，物理服务器的数量也随之减少【49】，从而服务器的能耗、制冷电器的用电量等大大降低，有利于绿色环保的企业环境的创建，响应国家关于节能减排的号召。6)解决平台的依赖性问题虚拟化技术的额出现使得我们将异构的操作系统进行整合，使得软件对运行平台的依赖性大大降低，能够满足业务在不同的平台架构、系统环境下运行150】【51】，满足同步发展与无缝连接，现在灵活共享。

云平台就是要把用户、软件、网络、计算、存储、数据库等完整的整合管理起来，并提供接口给上层使用者。其中用户和软件是使用资源的主题，网络、计算、存储、数据库都是资源，各用户按需动态获取资源来配置和调整自己的系统和服务，同时，系统还包括对用户的计费管理，用户软件程序的部署等。

织架构定义是云平台的基础，云管理平台需要支持多级组织嵌套，每级组织都会划分资源(CPU、内存、存储、网络等)和用户，并提供了多种用户的角色，各用户的功能视角也不同。根据数据中心云业务划分需求，云管理平台将使用者分为三大类：云管理员、租户管理员、最终用户。1)云管理员云管理员负责云数据中心的运维与运营，一方面云管理员需要负责云数据中心基础设施的运维工作，另一方面要负担起租户管理、流程管理、计费管理等运营工作。在大型的云数据中心中，云运维管理员和云运营管理员将会有不同人员负责。2)租户管理员租户管理员负责维护租户云资源，租户管理员可以对组织内云主机、云防火墙、云硬盘、云防火墙、云负载均衡、云数据库等进行管理。3)最终用户最终用户作为云资源的使用者，通过自助服务f-jp申请、使用、管理云资源，最终用户存在于租户之内。

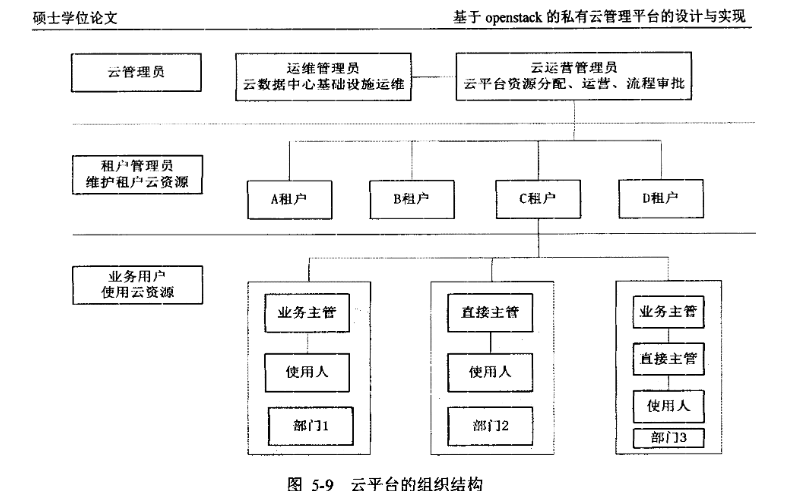


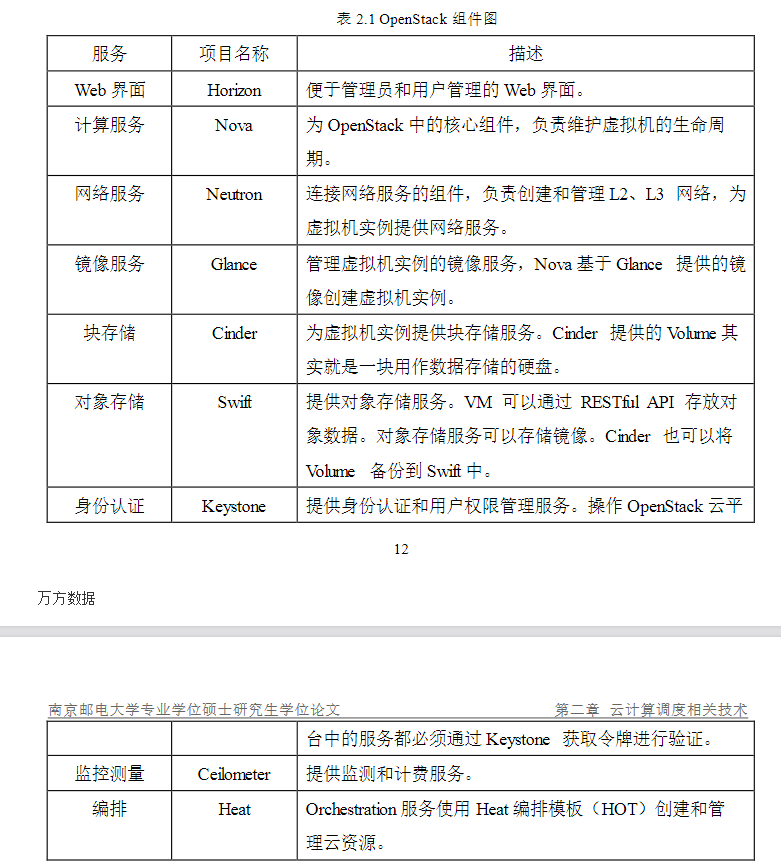


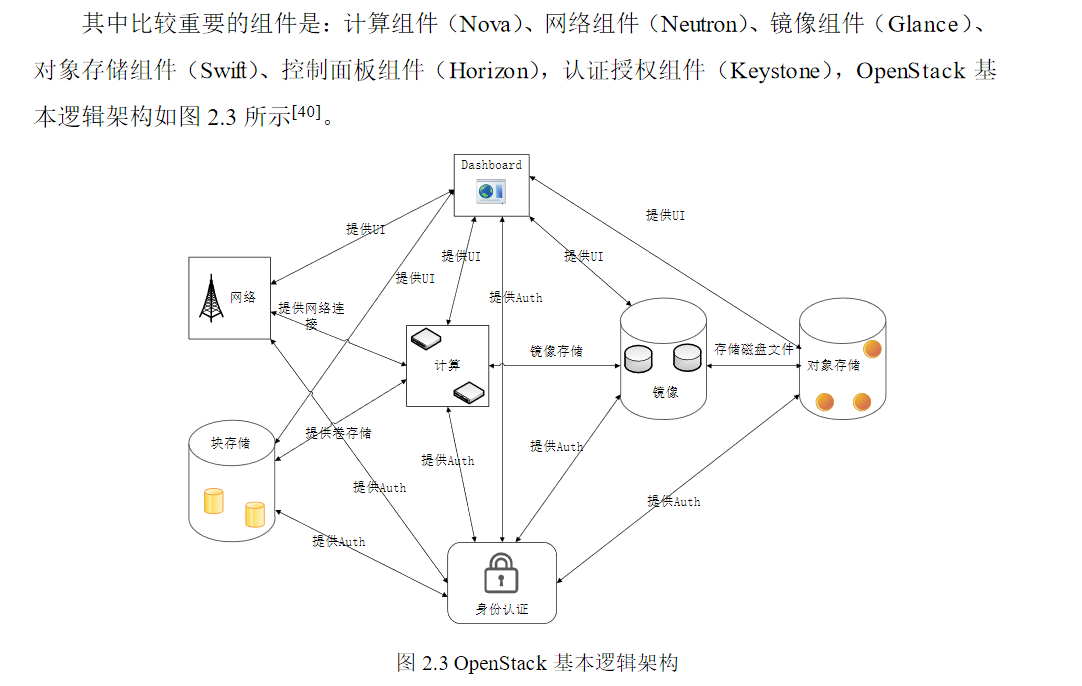
1. 云管理员云管理员对私有云的所有虚拟主机资源进行管理，包括创建主机、启动主机、访问主机控制台、关闭主机、重启主机、修改主机信息、销毁主机、查看主机列表、查看主机详细信息、创建主机规格、查看主机规格列表和查看主机规格详细信息。创建主机需要提供主机的规格(CPU、内存、硬盘等参数)，需要提供镜像模板(依赖于镜像模板模块)，创建主机前需要创建网络(依赖于网络模块)。启动主机后可访问主机控制台。查看主机列表后给出查看各个主机详细信息的链接，查看主机规格列表后给出查看各个主机规格详细信息的链接。2)组织管理员组织管理员对系统资源管理类模块的功能需求与云管理员大致相同，所不同的是管理范围。云管理员可管理私有云中所有虚拟化资源，组织管理员管理的是其所属组织和该组织的各下级组织的虚拟化资源。3)普通用户普通用户管理其所拥有的虚拟主机，包括申请主机、启动主机、关闭主机、重启主机、修改主机、销毁主机、查看主机列表、查看主机详细信息。

3．2非功能性需求3．2．1可扩展性需求为了实现大规模资源管理，同时支持业界标准，云平台基于Openstack的标准框架开发。通过定制Nova(计算)、Cinder(块存储)、Neutron(网络)、Glance(镜像)和Swift(对象存储)，提供面向接口与不同厂商的hypervisor和基础设施对接，实现对IaaS层资源的管理。在Keystone的基础之上丰富了租户的组织结构关系，满足政府和企业的复杂的行政关系。基于通用框架，智能云平台制定了标准，如基础架构标准、应用蓝本标准、业务迁移的标准。这样标准的架构可以同时满足私有云、公有云和混合云的模式，无论在任何场景中，数据和业务都可以灵活的迁移并应用于第三方标准平台。云平台的功能都提供业界标准的RESTAPI，标准开放的API使得云间资源的共享成为可能，多个云平台之间通过REST互访可以实现云间的资源灾备和迁移，完全解决了传统业务的信息孤岛问题。3．2．2性能需求1)查询时间：200个用户并发响应时间应当在5秒以内，一些大数据量(万条数据以上)的特殊数据(例如操作日志、工单、流程)的查询耗时应当在15秒以内。

用户管理模块主要功能包括用户的增加／申请、信息修改与查询、角色注销、角色管理、用户组管理等。当对用户进行上述操作后，系统通过用Email等方式通知用户操作结果。其中用户组管理是指对用户按照级别进行分层次维护和管理的功能。用户组管理实现对云管理平台中的用户分类并分层管理，运营平台支持将操作员划分为多个用户组，不同用户组的用户拥有不同的资源权限。角色管理是指权限的集合分类。用户可以将权限添加到角色中，用于方便权限的分配。整个系统的组织架构如图5-9所示

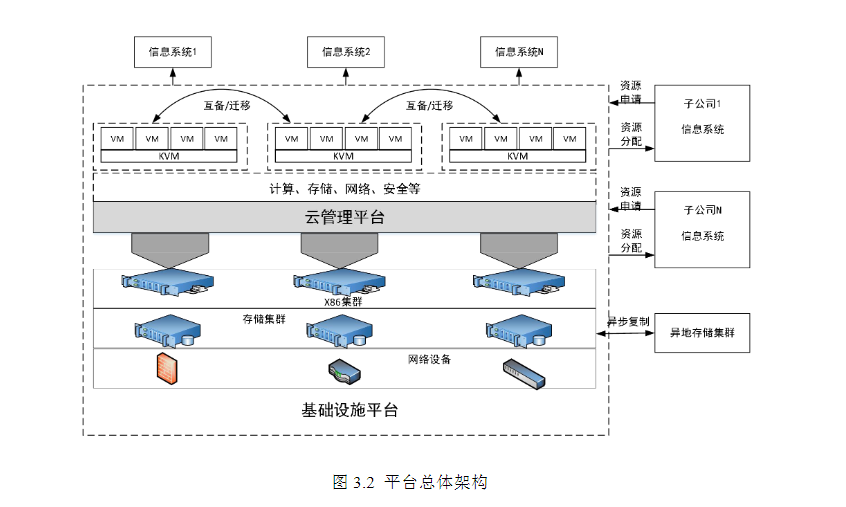


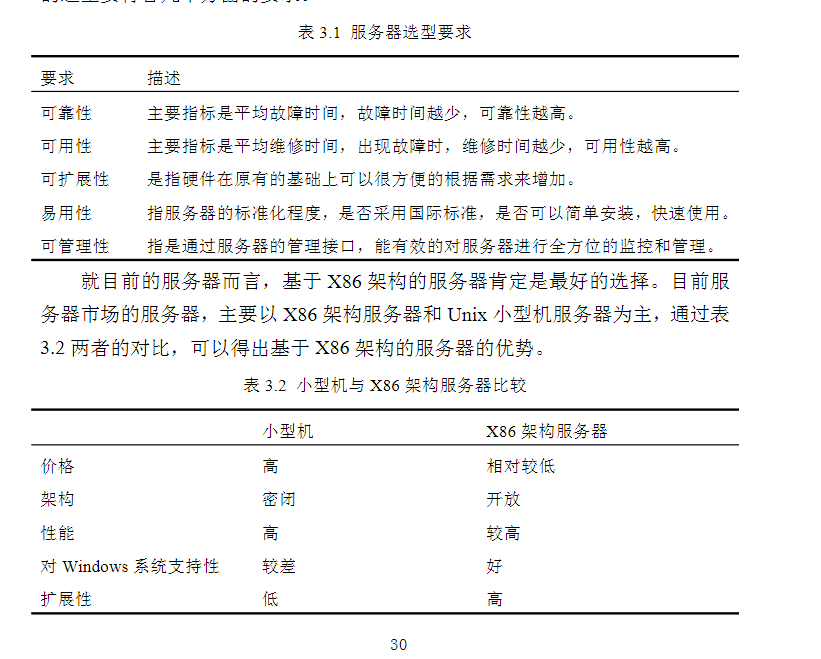




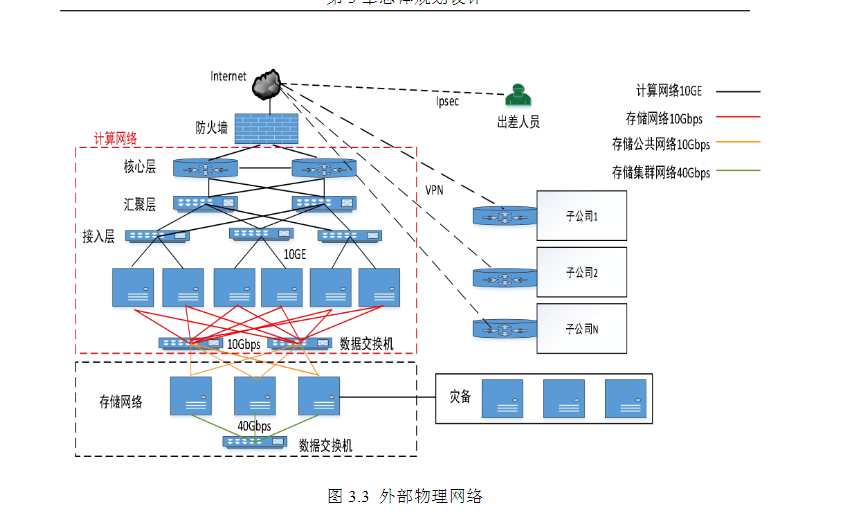
基于OpenStack私有云建设研究与应用

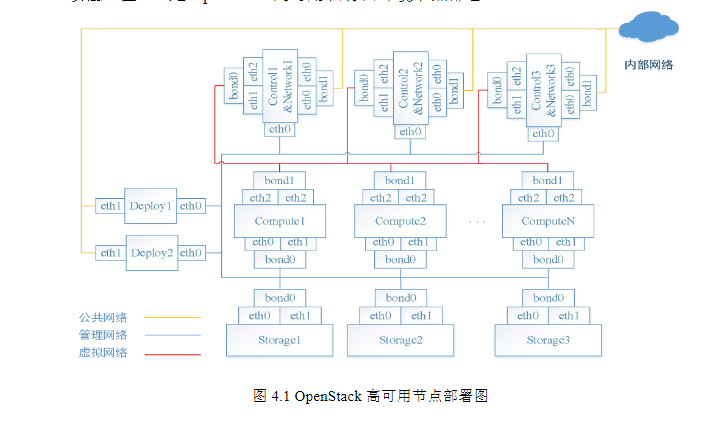
第3章总体规划设计283、循序渐进先建设一个小规模集群，核心业务暂不考上云，后续根据使用情况，可以利用云计算的优秀的扩展性，逐步增大规模。4、统一规划和分步实施私有云的建设是一个长期的系统的工程，绝非一朝一夕能够完成。遵循统一规划、分步实施的原则，首先实现IaaS层应用，后续逐步实现PaaS和SaaS层。3.3.2建设思路A公司数据中心的建设思路是：利用主流的开源技术搭建IaaS层面的私有云环境，充分利用云计算技术特点，建立一个资源动态分配，易扩展的云平台，来应对公司快速增长的业务需求。首先，通过虚拟化技术，将物理设备整合为统一的计算资源池、存储资源池和网络资源池[33]。在此基础上，通过云管理平台对资源进行动态调配，灵活扩展，实现硬件资源和软件资源的统一管理、统一运维、统一分配、统一备份，让IT基础资源更好的支撑应用系统的运行。基础设施平台为各子公司提供IaaS服务，各子公司不需要单独建数据中心和购买物理服务器等设备，可通过向总公司申请资源的方式，保障信息系统的正常运行，降低信息化的运营成本和管理难度。另外，为了保障业数据的安全性，可以在异地配置备份设备，并通过异步复制实现数据备份。3.4总体架构设计图3.2是公司数据中心的总体架构。根据需求，公司的数据中心通过私有云的建设，要实现统一管理、统一运维、统一分配、统一备份的IT基础设施环境，对各子公司提供IaaS服务，各子公司不需要单独建设数据中心和购买服务器等设备。要解决IT资源分散、需求快速增加；资源利用率低、服务器老化、耗能高且占用大量机房空间；数据无法统一备份、存在I/O瓶颈，异构环境难以整合；业务上线周期长、应用系统连续性无法保障；管理和运维难度大；建设成本居高不下等痛点问题。最终的目的是通过私有云的建设，降低信息化成本、提高资源利用率、实现资源动态调配、实现统一的IT基础设施环境，以便于支撑A公司快速的业务发展需求，为深化数据应用以及引入新技术、持续推动业务与管理高效打下基础。万方数据





公司的应用系统基本上都是基于Windows系统开发的，并且基本没有特殊的应用，所以公司私有云上采用X86架构的服务器更为合适。基于X86架构的服务器有机架式和刀片式的两类。刀片式服务器集成度高，但在扩展性和散热性上有明显的短板，并且后续还要为在机箱、后背板、交换模块上继续投入资金，网络容易被交换模块限制。机架式服务器具备灵活的扩展性，可以灵活的添加PCI-E模块和进行网络配置。而机架式服务器主要有二路和四路，二路多用于一般对性能要求不高的应用，四路多应用于高性能应用如数据库。综上所述，公司私有云建设采用X86架构的机架式服务器，并多以二路机架式为主，四路为辅。3、虚拟化软件选型目前主流虚拟化软件有：VMwareESXi、MicroSoftHyper-V、KVM、Xen等。但KVM和Linux内核高度集成，可以在内核内进行部署，很容易控制虚拟化进程。并且KVM更加灵活，它由操作系统直接和整合到Linux内核中的虚拟化管理程序交互，所以在任何情况下虚拟机都是直接与底层的硬件进行交互，不用修改虚拟机的系统。随着硬件对虚拟化的加速支持，为了提升性能和支持各种发行版本的操作系统，越来越多的用户转向了全虚拟化的KVM，如AWS和阿里云都在将底层Hypervisor由Xen迁移到KVM。另外，KVM是OpenStack默认的虚拟化引擎，结合OpenStack部署，基本不需要配置，能够有效的降低云平台的部署难度。综上所述，公司私有云虚拟化系统选择使用开源的KVM。





高可用是云平台必须考虑的重要问题，由于资源被整合在一起，一旦出现故障将影响整个平台的运行，对公司的正常生产和办公产生无法估量的影响。高可用[34]是指提供在本地系统单个组件故障情况下，能继续访问应用的能力，无论这个故障是业务流程，物理设施或是IT软硬件的故障。在集群服务方面，可将OpenStack高可用集群化分为基础架构服务高可用（包括消息队列高可用、数据库高可用和缓存服务高可用）、控制服务高可用、网络服务高可用、存储服务高可用和计算服务高可用。实现高可用的方法有很多种，目前，最常用的是基于Keepalived和HAProxy的实现方式。

4.5.1.1控制服务、数据库和消息队列高可用

为了满足控制服务、数据库和消息队列高可用，设计部署三台服务器做控制节点，具体实现如图4.5所示。OpenStack集群对外服务地址抽象层为一个虚拟IP地址，该IP的地址在Keepalived控制下实现控制节点之间的漂移，从而实现对外服务IP地址的高可用性。数据库以Galera实现高可用，采用MySQL和Galera组成三个Active节点，外部请求通过被Haproxy进行负载均衡后转发，万方数据

第4章基于OpenStack私有云设计与部署51转发以轮询的方式进行，一个控制节点出现故障，则会在剩余的控制节点上轮询转发，从而实现高可用。消息队列以同步镜像的方式实现高可用，它的实现原理与数据库的原理一样。

4.5.1.2网络服务高可用图4.6L3高可用方案考虑到目前和未来很长一段时间内，公司的集群规模都不是很大，L3高可用方案就可以满足网络高可用，图4.6为L3高可用实现方案。在三个控制节点上同时部署L3Agent服务，当创建高可用L3Router时，同时运行L3Agent的多个网络节点上同步创建多个L3Agent实例，不同网络节点上的L3Agent完全相同，借助VRRP使得多个网络节点上的L3Router具有不同万方数据

第4章基于OpenStack私有云设计与部署52的运行状态，仅有一个L3Router是Master状态，而其它的都是Standby状态。由Master享虚拟机提供路由服务，当Master出现故障时，会重新从Standb中选举新的Master，从而实现L3高可用。4.5.1.3存储服务高可用由于使用Ceph分布式存储系统，后端的高可用就可以由Ceph集群通过三备份的方式解决。前端高可用由HAProxy实现，HAProxy进行负载均衡后转发，转发以轮询的方式进行，一个控制节点出现故障，则会在剩余的控制节点上轮询转发，从而实现高可用。具体实现方案如图4.7所示。4.5.1.4计算服务高可用计算服务是OpenStack私有云中最核心的服务，目前最主流的计算服务高可用实现方案是Pacemaker\_remote。但它只能实现16个节点的高可用，并且对于虚拟机高可用来说，Pacemaker\_remote方案很有难度。这里的设计方案是：使用Docker容器化部署OpenStack，因为采用容器化部署，很容易发现计算节点故障。然后对计算节点进行周期性监测，发现故障后，通过Nova的EvacuateAPI进行虚拟机撤离，从而实现计算服务高可用。

根据公司的性质和情况，采用数据级的容灾方案，就能满足公司的需求，具体设计方案如图4.8所示。由于使用Ceph分布式存储系统，可以使用Ceph的RBDMirroring异步备份技术，完成数据备份。在外地的子公司中，使用大容量的SATA磁盘，组建一个Backup的CephCluster，利用RBDMirroring异步备份技术，完成异地数据备份。

4.5.3备份与恢复4.5.3.1数据库备份与恢复因OpenStack集群中的每个项目服务均使用相同的数据库引擎，但每个项目都有自己专属的数据库来存储数据，而这些数据库都位于同一个数据引擎中，所以可以很方便的实现数据备份。如要备份全部数据，可通过如下命令实现：

4.5.3.2配置文件与恢复OpenStack服务在启动运行时会读取配置文件信息，不同的配置文件参数将使得服务运行在不同的状态下，错误的配置文件和配置文件的丢失是致命的。不同的项目有不同的配置文件，需要各自备份。各个节点的要进行备份的文件分别是/etc/$project\_name、/var/lib/$project\_name和var/log/$project\_name。此外，/var/lib/docker/volumes和/etc/kolla这两个文件对于Docker容器化部署是很重要的，必须要做备份处理。



