十年一日，深入成就深度
业精于专，专注成就专业

数字有机体虚拟机系统

用户手册（V-1.20）

成都天心悦高科技发展有限公司

2016 年 2 月

版权声明

数字有机体虚拟机系统及其附属产品的版权属于成都天心悦高科技发展有限公司所有。任何组织和个人未经成都天心悦高科技发展有限公司许可与授权，不得擅自复制、更改该软件的内容及其产品包装。

本软件受版权法和国际条约的保护。如未经授权而擅自复制或传播本程序（或其中任何部分），将受到严厉的刑事及民事制裁，并将在法律许可的范围内受到最大可能的起诉！

成都天心悦高科技发展有限公司
地址：成都市武侯区棕南小区
电话：028－83318559
邮编：610054

目录

1. 引言

1.1 编写约定

非常感谢您使用成都天心悦高科技发展有限公司的产品，本公司将竭诚为您提供最好的服务。

本手册可能包含技术上不准确的地方或文字错误。

本手册的内容将做定期的更新，恕不另行通知；更新的内容将会在本手册的新版本中加入。

本公司随时会改进或更新本手册中描述的产品或程序。

1.2 内容简介

本文档供数字有机体虚拟机系统的系统管理员和项目管理员阅读，帮助他们使用数字有机体虚拟机系统，实现存储、网络和计算等的虚拟化，运行各种虚拟机。

本文档分为 10 个章节。第一章即本章。第二章初步介绍数字有机体虚拟机系统，如果需要了解本系统的功能，请参考这章。第三章讨论虚拟化系统的各种可能运行方案，帮助系统管理员规划存储、网络和计算资源的虚拟化方案，并构建合适的部署方案。第四章介绍系统软件的安装。从第五章开始介绍系统的操作界面和功能。本文将系统管理员和项目管理员的操作界面放在同一个章节中介绍，原因是大多数功能需要两者协作完成。第五章介绍用户和个人信息管理。第 6 章介绍系统服务器管理。第 7 章介绍作为运营管理的基础——项目的管理。第 8 章介绍虚拟机运行需要的安全对象的管理。第 9 章讨论存储虚拟化管理，包括如何建立虚拟机运行需要的镜像文件（光盘或者磁盘的映像）和存储卷。第 10 章讨论网络的虚拟化，包括子网管理，网络过滤器和网络管理。第 11 章介绍虚拟机运行的管理工作。第 12 章介绍数字有机体虚拟机系统特有的功能，虚拟机集群的管理。第 13 章介绍其他的辅助功能，包括操作日志审计和系统参数设置。最后一章为出错处理，告诉用户在系统出现问题时怎样解决。

1.3 相关文档说明

数字有机体虚拟机系统基于数字有机体平台，有关数字有机体平台的文档如下：

- 有关数字有机体系统的安装，请参阅《数字有机体系统安装指南》的描述。
- 有关数字有机体工作平台的使用, 请参阅《数字有机体工作平台及抗毁容灾系统用户手册》。
- 有关如何在数字有机体工作平台上开发应用程序，请参考《数字有机体工作平台及抗毁容灾系统开发手册》。
- 有关数字有机体工作库的使用, 请参阅《数字有机体大规模存储与管理系统用户手

册》。

- 有关数字有机体文件系统的使用，请参阅《数字有机体工作平台及抗毁容灾系统用户手册》。

1.4术语

虚拟机资源池：某个子网内的相关成员的集合。并对选定的子网内相关成员按选定负载均衡方式执行负载均衡。

LVM：Logical Volume Manager， 逻辑卷管理器。它是 Linux 环境下对磁盘分区进行管理的一种机制，LVM 是建立在硬盘和分区之上的一个逻辑层，来提高磁盘分区管理的灵活性。通过 LVM 可以轻松管理磁盘分区，如：将若干个磁盘分区连接为一个整块的卷组，然后在卷组上创建卷，并在卷上建立文件系统。要注意，LVM 的卷和存储虚拟化中的卷不是一个概念。

RAID：独立磁盘冗余阵列，用多块磁盘构建的磁盘系统，以便扩展容量，提升读写性能，或者增强数据的可靠性。

存储池（storage pool）：一些存储设备在软件的管理下形成的可以共享的存储空间。在虚拟机系统中这是一个虚拟的概念，它表达了以下几个意思：1）存储池是共享的，可以将其中的空间分配给不同的虚拟机使用。2）存储池是可以管理的，管理员可以在其中创建卷，改变卷大小等。3）存储池只是一个逻辑概念，它可以采用不同的方式实现，形成不同类型的存储池，例如文件目录存储池、IP-SAN 存储池、磁盘卷存储池等。

存储卷（volume）：它是存储池中建立的可独立分配的存储单元。它可以映射为各种具体的实体。具体是那种实体由卷所在的存储池决定。

镜像文件（image）：一种有格式的文件，用以模拟某个存储设备。镜像文件需要由具体的工具来支持。在本系统中，镜像文件由 qemu-img 处理。qemu 是一个著名的虚拟机模拟器。它支持很多种镜像文件格式，用以模拟光盘（iso9660）、软盘和硬盘(qcow、qcow2、raw 等)。

iSCSI：它是一个供硬件设备使用的可以在 IP 协议的上层运行的 SCSI 指令集，这种指令集可以实现在 IP 网络上运行 SCSI 协议，使其能够在诸如高速先兆以太网上进行路由选择。iSCSI 是一种新存储技术，该技术是将现有的 SCSI 接口与以太网络技术结合，使服务器可与使用 IP 网络的存储装置互相交换资料。

SAN：它是 Storage Area Network 的缩写，意为存储区域网络，是一种网络存储系统，用于实现海量存储的集中部署和网络访问。根据传输网络的不同，可以分为 FC-SAN 和 IP-SAN。

FC-SAN：光纤存储区域网络，用光纤网络互连的存储区域网络。

IP-SAN：基于 IP 网络的存储区域网络，其优势是可以利用高速以太网的成本优势。在本系统中，它提供的存储池成为 iSCSI 存储池。

NAS：它是 Network Attached Storage 的缩写，意为网络附属存储，按字面简单说就是连接在网络上，具备资料存储功能的装置，因此也被称为“网络存储器”。它是一种专用数据存储服务器。它以数据为中心，将存储设备与服务器彻底分离，集中管理数据，从而释放

带宽、提高性能。目前著名的 NAS 企业有 Netapp、EMC、OUO 等。

NFS：Network File System, 网络文件系统的缩写。NFS 通常作为一种网络文件服务使用，客户通过网络文件系统访问集中存储和管理的文件，从而实现数据共享和存储共享。

CIFS：它是一种协议，它使程序可以访问远程 internet 计算机上的文件并要求此计算机提供服务。CIFS 使用客户/服务器模式。客户程序请求远在服务器上的服务器程序为它提供服务。服务器获得请求并作出响应。CIFS 是公开的或开放的 SMB 协议版本，并由 Microsoft 使用。SMB 协议在局域网上用于服务器文件访问和打印的协议。像 SMB 协议一样，CIFS 在高层运行，而不像 TCP/IP 协议那样运行在底层。CIFS 可以看做是应用程序协议如文件传输协议和超文本传输协议的一个实现。

桥接（bridge）：用网桥方式连接虚拟网络接口和物理网络的方式。这里特指在操作系统中实现的虚拟的网桥设备，然后通过虚拟网桥在虚拟网络接口和物理网络接口间实现桥接。可以认为网桥就是老式的 HUB，只是速度更快。

Macvtap：一种由 Linux 内核实现的虚拟网络设备，它直接在物理设备和虚拟设备间交换数据，即从物理设备进来的数据将在各个虚拟设备上发出，从虚拟设备进来的数据则从物理设备上发出。这使得虚拟设备间不能相互通信。

OVS：Open Virtual Switch，开放的虚拟交换机。一个用软件实现的虚拟交换技术，用于模拟路由交换机。

Open-VSwitch：一个开源的虚拟网络组建软件。

DHCP：它是 Dynamic Host Configuration Protocol 的缩写，意为动态主机配置协议，是一个局域网的网络协议，使用 UDP 协议工作，主要有两个用途：给内部网络或网络服务供应商自动分配 IP 地址，给用户或内部网络管理员作为对所有计算机作中央管理的手段，在 RFC2131 中有详细的描述。DHCP 有 3 个端口，其中 UDP67 和 UDP68 为正常的 DHCP 服务端口，分别作为 DHCP server 和 DHCP client 的服务端口，546 号端口作用于 DHCPv6 client，而不用于 DHCPv4，是为 DHCP failover 服务，这是需要特别开启的服务，DHCP failover 是用来做“双机热备”的。

DNS：它是 Domain Name System 的缩写，意为域名系统，因特网上作为域名和 IP 地址相互映射的一个分布式数据库，能够使用户更方便的访问互联网，而不用区记住能够被机器直接读取的 IP 数串。

KVM：它是 Kernel-based Virtual Machine 的缩写，意为基于内核的虚拟机，是一个开源的系统虚拟化模块。KVM 的虚拟化需要硬件支持。是基于硬件的完全虚拟化。

NAT：它是 Network Address Translation 的缩写，意为网络地址转换，是 1994 年提出的。当在专网内部的一些主机本来已经分配到了本地 IP 地址（即仅供本专网内使用的地址），但现在又想和因特网上的主机通信时，可以使用 NAT 的方法。另外，它还可以解决使用少量的公网 IP 地址代表较多的私有 IP 地址，将有助于减缓可用的 IP 地址空间枯竭。在 RFC1632 中有对 NAT 的说明。

IP 隧道：IP 隧道技术是路由器把一种网络层协议封装到另一个协议中以跨过网络传送到另一个路由器的处理过程。

VLAN：它是 Virtual Local Area Network 的缩写，意为虚拟局域网。它是一组逻辑上的

设备和用户，这些设备和用户并不受物理位置上的限制，可以根据功能、部门及应用等因素将它们组织起来，相互之间的通信就好像它们在同一个网段中一样，由此得名虚拟局域网。

VIP：它是 Virtual IP 的缩写。虚拟机集群通常用于向网络用户提供某种服务。这个服务需要关联一些具体的参数，例如服务 IP 地址、服务端口、最大连接数、采用的网络协议和是否要求保持会话等。本系统将服务的这些参数视为一个逻辑实体，称作 VIP。

虚拟机实例：包含完整虚拟机定义的虚拟机配置称作实例。

虚拟机： 根据虚拟机实例的属性和行为模拟出来的一台虚拟的计算机。

虚拟机资源池：又称为虚拟机集群，某个虚拟网络内的相关虚拟机实例（及由实例模拟出来的虚拟机）的集合。系统为这些虚拟机提供网络负载均衡支持，并可按照规则自动增减虚拟机数量，从而形成一个服务资源池。

安全组：安全组是 IP 过滤规则的集合，包括防火墙控制等，用于控制虚拟机的网络安全访问管理，比如对于公网运营的虚拟机 Web 服务器，一般开通 TCP/80 端口用于页面访问。安全组由很多网络访问规则组成。

过滤器：安全组中设定的网络访问规则被称为过滤器。需要按照 Libvirtd 的描述语法定义。它支持一个过滤器调用另一个过滤器。

宿主机群：由于虚拟机组有相同的网络需求，因此可以为其指定一组用以运行虚拟机的宿主机。这组宿主机称为宿主机群。

1.5 如何获得技术支持

在您遇到问题时，请首先联系您的产品提供商。大多数问题都可以在产品提供商的技术支持人员的帮助下得以解决。

您也可以通过产品提供商致电本公司的技术服务热线：028-83318559，获得电话技术支持。您还可以发送邮件，邮件地址是：tianxinyue@126.com。如果您确实需要本公司提供上门服务，本公司将竭诚为您服务。

2. 简介

2.1 数字有机体系统简介

数字有机体系统（英文名称为 Digital Organism System，缩写为 DOS）是在刘心松教授带领下，由成都天心悦高科技发展有限公司的研发人员前后千余人次，经过三十多年的技术积累，研发成功的基础系统。

研发这种系统的原始宗旨是向生物特别是人类个体和群体的结构、机理和特性逼近，是一种人能化的新的系统模式。这种系统集操作系统、数据库系统、大规模存储、抗毁容灾、高伸缩、高智能、高灵活、自搜索、自传播、自复制、自修复、自重构、自适应、系统间的兼容性、群体间的协作性、对资源的动态管理调度合理配置、大小新旧机器混合使用等特性为一体，是一个整体解决方案，是面向所有应用的统一的（应用）系统平台。

数字有机体系统主要由数字有机体工作平台、数字有机体抗毁容灾系统、数字有机体工作库、数字有机体大规模存储与管理系统、数字有机体安全系统组成。这是从底层作起的一个一体化平台，可以在此平台上开发任何应用，形成任何应用系统。例如现在已有的应用系统就有数字有机体流媒体系统、数字有机体监控系统、数字有机体会议系统、数字有机体网关、数字有机体管理系统、数字有机体控申系统、数字有机体侦查指挥系统等。

本文有时将数字有机体工作平台及抗毁容灾系统，数字有机体工作库及大规模存储与管理系统和数字有机体安全系统统称为数字有机体系统。数字有机体工作平台及抗毁容灾系统含盖常规操作系统但远高于常规操作系统，是一个在 Linux 之上的、面向很多应用的、统一的、人能化的应用系统平台。数字有机体工作库及大规模存储与管理系统含盖常规数据库系统但远高于常规数据库系统，是一个在 Mysql 之上的、面向很多应用的、统一的、人能化的应用数据平台。

有时将数字有机体工作平台及抗毁容灾系统简称为数字有机体工作平台甚至工作平台。

有时将数字有机体工作库及大规模存储与管理系统简称为数字有机体工作库甚至工作库。

2.2 数字有机体虚拟机系统简介

从整合应用、方便管理和充分利用资源等需求出发，数据中心不断向云计算模式迁移。而云计算的核心是资源的虚拟化，即存储虚拟化、网络虚拟化和计算虚拟化。数字有机体系统能够同时实现上述三者的虚拟化，这就是数字有机体虚拟机系统。

和其他的云计算平台相比，数字有机体虚拟机系统具有以下特色：

1）系统不依赖于大型存储设备、大中型计算机或者专用的网络设备。数字有机体虚拟机系统希望充分利用高性价比的服务器、磁盘阵列或者内置磁盘，以及高性价比的以太网络。通过整合这些高性价比的资源，构建可靠的、高性能的、资源利用率高的系统。

2）系统没有中心节点或者固定的控制节点，各台服务器分工协作，在逻辑上都是平等

的。即使是网络虚拟化，也不建议采用集中的网络节点。这样的系统具有更高的可靠性，且更能适应变化的环境。

从功能上讲，数字有机体虚拟机系统的核心功能是三化加管理，即计算虚拟化、存储虚拟化、网络虚拟化，加集成管理。

计算虚拟化以虚拟机模拟器（monitor，亦称监视器）为基础，通过虚拟机动态调度、定时启停虚拟机、自动增减虚拟机以及虚拟机网络负载均衡，达到将系统资源动态地按需分配给应用的目的。这里的按需分配是有确切功能支撑的。大多数云计算平台只是通过人工部署或者启停虚拟机来达到调整各个业务占有资源的目的。在数字有机体虚拟化系统中，系统可以自动地按照用户设定，定时启动和停止虚拟机的运行；甚至可以自动地根据应用当前的负载情况增加和减少虚拟机数量，从而达到完全动态的自动增减应用资源的目的。配合虚拟机动态运行调度，可更有效的利用系统的各种计算和网络资源等。

数字有机体虚拟机系统的存储虚拟化可以通过多种途径实现。如果喜欢集中部署的存储系统，可以采用 FC-SAN 或者 IP-SAN。这时存储虚拟化仅仅在 SAN 系统中实现，虚拟机使用固定的由 SAN 提供的逻辑单元（LUN）。虚拟机可以运行在能够访问 SAN 系统的服务器上。这种部署方式的优点是集中，缺点是集中导致的部署限制、性能限制、容量限制和可用性限制。

另一个可供选择的存储虚拟化方案是数字有机体虚拟存储。它支持存储设备分散部署，支持异构的存储设备，支持广域网络分散部署，支持大量虚拟机并发访问。在数字有机体系统中，存储设备连接到服务器上，由服务器直接管理，再由数字有机体系统进行整合。数字有机体虚拟存储系统的其主要功能是：

1）整合分散部署在各地的服务器上的存储资源形成统一的虚拟存储池。

2）将这个存储池中的资源按需分配给虚拟机使用。

3）为虚拟机模拟器提供位置无关的访问服务，使虚拟机可以运行在系统中任何一台服务器内。

和其他云平台相比，数字有机体存储虚拟化的不同点在第一个功能上。数字有机体能聚合分散存储资源构建存储池，而不是直接利用物理存储系统，如 FC-SAN 或者 IP-SAN 这样的大型存储系统做存储池。

网络虚拟化的核心是为虚拟机运行提供网络支撑。数字有机体虚拟机系统的网络虚拟化功能是：1）为虚拟机通信提供虚拟设备。2）构建虚拟机通信需要的虚拟网络。3）为虚拟机通信提供安全隔离。4）为虚拟机通信提供网络服务，如 MAC 分配、IP 地址分配、交换服务等。和 Open Stack 相比，数字有机体虚拟化系统构建的是分布式的通信网络，而不是由集中节点控制的网络。

在核心的虚拟化功能之上，数字有机体虚拟机系统也具有完善的管理系统。这个管理系统将和底层功能配合，为管理员控制和监控系统，应用部署人员部署和运行应用的虚拟机，终端用户使用虚拟机提供各种方便快捷的操作终端。当然，还包括安全控制、运行审计等功能。

2.3 主要功能和特点

2.3.1 主要的功能

数字有机体虚拟机系统主要用于部署和管理私有云。它具有以下功能：

1) 分布式部署：支持大量服务器通过高速网络互联，形成统一的系统，用于部署各种虚拟机，且支持多地分散部署。
2) 多种操作系统的虚拟机：支持部署多种操作系统的虚拟机，例如 Linux 和 Windows 系统。
3) 按需分配：各个虚拟机实例根据用户的配置需求分配 CPU 核数、CPU 类型和内存大小。
4) 多虚拟机运行：同一台宿主机上可以运行不同的虚拟机实例，每个虚拟机实例可以同时运行很多个具体的虚拟机。
5) 定时自动启停：可以根据管理员的配置，在某个时间段自动启动或者停止某个或者某些虚拟机实例，以满足某个业务随时间波动的资源需求。
6) 自动增减虚拟机：系统自动探测同一虚拟机实例的所有虚拟机的负载情况，如果都超过某个限值则自动增加虚拟机，如果都低于某个限值则自动减少虚拟机，从而实现按需动态分配资源的目的。
7) 虚拟机运行位置控制：系统可以根据设定的要求（如绑定宿主机或者仅在某个宿主机群中运行）和各台服务器的负载情况自动选择虚拟机的运行位置，也可以通过管理系统指定虚拟机的运行位置，还可以在线迁移虚拟机。
8) 负载均衡：当虚拟机运行位置可自动选择时，系统根据各台服务器的能力和负载情况选择虚拟机的运行位置，从而均衡各台服务器的负载。
9) 多虚拟机整合：系统支持虚拟机集群，提供网络负载均衡功能以使多个虚拟机可以共同提供某项服务。当这些虚拟机实例运行在不同的宿主机上时，就达到了整合多台宿主机服务能力的目的。即使这些虚拟机都运行在同一个宿主机内，也可以通过自动增减虚拟机达到按照业务负载情况动态分配资源的目的。
10) 支持多种类型的存储池：在单台服务器上，支持使用一个磁盘设备构建磁盘设备池，支持通过 LVM 实现逻辑卷池，支持通过本地文件系统实现文件目录池。支持利用 IP-SAN 实现 iSCSI 存储池。支持利用 NAS 或者文件服务系统构建网络文件系统池。也可以通过数字有机体系统构建高扩展的、高可靠的、支持无限容量的数字有机体文件系统池。
11) 磁盘设备虚拟服务：系统为虚拟机提供虚拟的磁盘设备。虚拟机可以根据自己的需要使用这些磁盘设备。每个虚拟机可以同时有多个磁盘设备。每个磁盘设备的大小可以不同。无论虚拟机在那台服务器上运行，都能访问到这些磁盘设备。
12) 按需配置磁盘设备：每个虚拟机实例可以根据需要配置磁盘设备的数量、每个磁盘设备的大小。
13) 虚拟磁盘设备扩容：可以根据用户的需要增加某个虚拟磁盘设备的大小，从而让用户获得更大的空间。

14) 存储的虚拟机实例无关性：用户可以根据自己的需要，将某个虚拟磁盘设备连接到另一个虚拟机实例，从而在另一个虚拟机实例上使用该磁盘设备。
15) 存储位置无关性：通过数字有机体文件系统存储池，或者其他网络存储池，使虚拟机可以在不同的宿主机上访问同一个存储卷，即虚拟机的磁盘。
16) 虚拟机通信支持：支持虚拟机通过 NAT、路由、桥接等方式和其他虚拟机或者外部网络通信。
17) 虚拟网络支持：支持通过 Linux 桥接器、OVS 桥接器和 Macvtap 等方式构建各种虚拟网络，自动部署网络网关等。
18) 网络服务：提供 DHCP、DNS 和 tftp 网络服务，方便虚拟机的网络配置和虚拟机运行等。
19) VLAN：通过 Open VSwitch 支持 VLAN，实现虚拟子网间的安全隔离。
20) 网络安全：支持配置网络安全组，用以保护虚拟网络及其中的虚拟机的网络安全。
21) 虚拟机图形化配置：在管理系统中，可以通过 WEB 界面轻松的配置虚拟机。
22) 虚拟机自动部署：在管理系统中配置好虚拟机后，可以通过管理系统直接在某台宿主机上启动虚拟机。系统自动完成虚拟机运行环境的配置，并使虚拟机运行起来。
23) 虚拟机迁移：支持在相同结构的宿主机间在线迁移虚拟机。
24) 虚拟机保存恢复：支持保存虚拟机的当前运行状态，在需要时再恢复运行。
25) 虚拟机暂停继续：支持暂停虚拟机，然后再继续。
26) 虚拟机快照：通过建立虚拟机检查点（即快照），可保存虚拟机在某个时刻的状态，在需要时可以回到某个快照继续运行。
27) 远程虚拟机桌面：支持通过 VNC 或者 Spice 远程访问虚拟机桌面。
28) 远程控制虚拟机：可以远程启动、关闭、重启或者重置虚拟机，或者为虚拟机建立快照，保存虚拟机等。
29) 远程监控：支持通过管理界面远程监控虚拟机的运行状况，包括虚拟机负载等。
30) 按项目组织虚拟机：当系统规模很大时，需要同时管理许多虚拟机。系统支持按照项目组织虚拟机。

2.3.2 系统特点

数字有机体虚拟机系统具有以下特点：

1) 兼容大小新旧服务器：系统由性能不同，新旧不同，品牌不同的服务器构成。即可能存在大型机，也可能有普通服务器。
2) 大规模分布式部署：系统可能由部署在几百上千的数据中心（或者部署点）的数字有机体站构成。系统具有大量服务器，而不是少数服务器。
3) 不依赖物理上集中的存储系统：存储系统由有各台服务器共享出的存储空间构成，而不是由单独的 IP-SAN 或者 FC-SAN 构成。重要的是，设计时不能假设有共享的存储设备可供使用。
4) 无中心节点：系统不能存在集中的固定的控制节点或者管理站点。
5) 无单一故障点：系统不会因单台服务器或者单个数据中心的故障而停止服务。

3. 部署方案规划

3.1 基本概念

1）什么是虚拟机？

虚拟机是和物理计算机相对应的。一台物理上的计算机是真实存在的实体，它通常由中央处理（CPU）、内存、磁盘设备、网络接口、系统总线，以及其他各种外部设备构成。在某些环境下，希望在一台真实的物理计算机上同时运行多个独立的任务。典型的，这些任务甚至需要不同的操作系统。这时，就需要将一台物理计算机虚拟为多台虚拟的计算机，这就形成了虚拟机。更近一步，虚拟机需要使用的存储资源和网络资源也可以网络化和虚拟化，从而形成了现在的虚拟机环境。

2）计算虚拟化

计算虚拟化是，通过虚拟化技术，将计算资源分解或者整合，实现按需分配和动态分配，以便更加充分的利用计算资源。对大中型计算机来说，大多数业务都无法充分利用这台计算机的资源。而将多个业务整合到一台计算机上运行时，面临各个业务需要的操作系统等软件环境可能不同，需要的资源数量可能不同的问题。解决这个问题的办法是虚拟机技术，即通过软件来模拟硬件，在一台物理计算机上模拟出多个虚拟的计算机，从而能够让不同的业务在不同的虚拟计算机中运行。

虚拟机技术解决了分解大型资源，按需分配给业务的问题。但是，如果不在中大型计算机上，而在普通的服务器时，则可能一台服务器无法满足业务的资源需求。这就需要整合计算资源。整合计算资源的主要技术是分布式并行技术，包括分布式计算、集群技术等。

计算资源主要指计算处理能力，对应的是 CPU 和内存，以及支持 CPU 运行的环境，如总线、缓存等。

3）存储虚拟化

存储虚拟化的需求和计算虚拟化类似。采用的方案有几种。第一种是由硬件实现的虚设备。最初是通过 RAID 技术实现的虚磁盘设备，再往后是 FC-SAN 提供的虚网络存储设备，以及现在流行的 IP-SAN 提供的虚逻辑单元。第二种是通过网络文件系统实现的存储虚拟化。如 NAS 系统以 NFS 和 CIFS 方式提供文件服务，集群文件系统提供的文件服务，分布式对象存储系统提供的对象存储服务等。数字有机体系统提供基于广域网络的共享文件系统，也可以属于这类。在文件系统的基础上，可用单个或者多个文件来虚拟一个磁盘，这就需要虚拟机的磁盘虚拟功能来完成，如 Qemu（一个虚拟机模拟器）支持的 QCOW、QCOW2、RAW 等格式的磁盘镜像文件。

由硬件实现的虚拟存储性能更好，稳定性也更好，但是成本高昂，扩展性不足，动态分配性差，且在抗毁容灾能力上不如基于文件系统的方式。基于文件系统的优点是构建成本更低，扩展性好，动态分配能力强，可屏蔽故障，可靠性更好。数字有机体系统更具有抗毁容灾、广域部署、动态扩展等特性，是良好的虚拟化存储解决方案。

4）网络虚拟化

网络虚拟化的需求来自多个方面。在虚拟机系统中，一方面是为虚拟机提供通信网络，这包括提供虚拟网络接口，提供主机内虚拟机间直接通信的手段，提供共享宿主机网络接口与实体网络通信的手段等。另一方面是虚拟机组网。在同一台宿主机内的虚拟机也可能需要划分为不同的子网，也可能需要路由、地址转换和 VLAN 等。这些功能由软件实现的虚拟路由交换设备完成，如 Open-VSwitch 等。如果系统由大量宿主机构建，就存在不同宿主机上的多个虚拟机要相互通信的问题。解决的办法是虚拟桥接器，如 Linux 的桥接器、Open-VSwitch 的桥接器等。桥接技术将虚拟机的虚拟接口直接桥接到物理网络上，这使得可以通过物理网络设备来构建虚拟机的网络。在虚拟机系统中还存在网络通信的其他需求，如流量控制、访问控制、通信过滤、QOS 等。这些也属于网络虚拟化的范畴。

除了虚拟机通信需要的网络虚拟化外，现在还在研究完全由软件实现的，可自由调整网络结构的组网技术，如 Open Flow 等。这使得物理网络本身也会虚拟起来。不过，虚拟机通信暂时还不是必须这样的虚拟化。

无论是计算虚拟化、存储虚拟化还是网络虚拟化，都存在许多不同的实现技术。这些技术能够虚拟的程度可能不同，要解决的问题也可能不同，具有的优缺点也不同。这里无法一一介绍这些技术。

3.2 存储虚拟化方案

在确定存储虚拟化方案前，请先收集如下信息：

1）需要使用存储空间的虚拟机的部署范围：单机、集中部署的多台服务器、集中部署的大量服务器、分散部署的服务器；

2）总存储空间需求；

3）存储设备共享需求：固定划分、动态分配；

4）数据共享需求：有或者无；

5）存储管理需求：集中或者分权管理。

数字有机体虚拟机系统支持多种存储虚拟化方案。典型的存储虚拟化方案有：

1）单机服务器存储虚拟化；

2）IP-SAN 存储虚拟化；

3）NAS 存储虚拟化；

4）数字有机体存储虚拟化。

3.2.1 单服务器存储虚拟化

如果虚拟机只部署在一台服务器上，系统的总存储空间需求不大，则可以采用这种部署方案。这时，用于运行虚拟机的存储设备可以是服务器内置的磁盘系统，也可以是服务器外接的磁盘阵列。虚拟机可以采用三种方式使用存储设备。

1）磁盘卷池

采用这种方式时，一个独立的磁盘就是一个存储池。外接的磁盘阵列通常也映射为服务器上的一个逻辑磁盘，因此外接的磁盘阵列也可用于构建磁盘卷池。在磁盘中，可划分多个

分区，每个分区就是一个卷。每个卷可以给一个运行的虚拟机使用。采用这种方式的优势是：各个虚拟机间的存储空间隔离性好，虚拟机的磁盘性能只受物理磁盘的影响，额外开销小。缺点是存储是固定分配的，无法实现动态分配，且无法整合多个磁盘的空间。

2）逻辑卷池

如果服务器同时有多个硬盘，或者在服务器上同时有多个逻辑设备，又希望将这些逻辑设备的存储空间整合起来，就可以使用逻辑卷池。这需要利用操作系统的 LVM（逻辑卷管理，LVM 的卷和本文谈到的卷是两个不同的概念）功能将多个逻辑设备整合为一个 LVM 卷。这个 LVM 卷可作为一个逻辑卷池。然后在 LVM 卷中创建多个 chunks，每个 chunks 就是一个存储卷。再将这些存储卷分配给虚拟机使用。

这种方式可以整合服务器上的多个逻辑存储设备，也便于给虚拟机分配卷。但是性能因 LVM 有很少量的降低。

3）宿主机本地目录池

如果需要动态的给虚拟机分配空间，且希望这些虚拟机共享存储设备，则可以采用宿主机本地目录池。这时，通常先采用 LVM 或者 RAID 技术整合逻辑设备（单个逻辑设备则无需整合），然后将逻辑设备安装（mount）到本地目录下，以文件系统（建议采用 EXT4 文件系统）的方式使用。将这个本地目录作为一个存储池（当然也可以在文件系统中建立多个目录，分别作为存储池），再在存储池中创建镜像(image)文件作为存储卷，供虚拟机使用。

采用这种方式时，存储卷的大小可以是不固定的，也无需预先分配存储空间，在虚拟机运行过程中再动态分配。缺点是存储的访问性能因多层处理而有部分损失。

单服务器部署的存储虚拟化方案只适合小规模的应用，无法支撑大型应用的运行。

3.2.2 IP-SAN 存储虚拟化

如果虚拟机需要部署在集中部署的多台服务器上，或者集中部署的大量服务器上，总存储空间需要很大，存储访问性能要求高，数据共享需求不大，且需要集中管理，则建议采用 IP-SAN 存储虚拟化方案。其部署如图 3-1 所示。

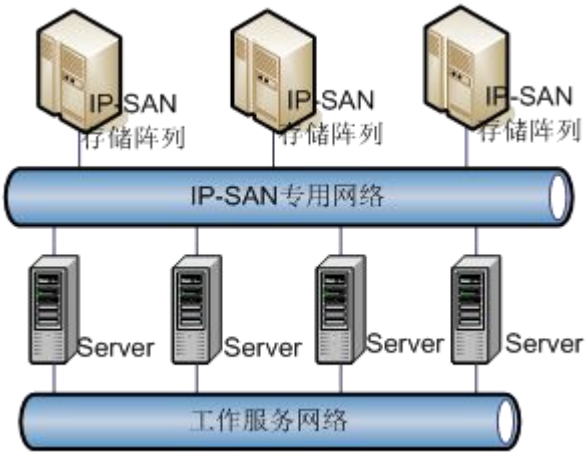


图 3- 1：IP-SAN 的存储虚拟示例

这时需要采用单独的 IP-SAN 存储系统作为存储资源。建议服务器采用独立的网络接口访问 IP-SAN 系统。在 IP-SAN 系统中，为每个虚拟机划分存储空间，即逻辑单元（LUN）。

可将整个 IP-SAN 系统作为一个存储池，也可以根据需要将 IP-SAN 划分为多个存储池，这些需要在 IP-SAN 系统中配置。对虚拟机来说，使用的存储卷就是 IP-SAN 系统的一个逻辑单元（LUN）。

采用IP-SAN存储虚拟化方案时，需要在数字有机体虚拟机系统的管理系统中建立iSCSI存储池，然后通过刷新的方式获得存储池中已经建立的卷（即 LUN）。数字有机体虚拟机系统还不支持建立 iSCSI 池中的卷。

采用这种方式的优势是：虚拟机存储设备的访问性能高。但缺点是无法大规模分散部署，存储空间预先分配，且可靠性受限于单一的 IP-SAN 系统。

3.2.3 NAS 存储虚拟化

NAS 存储系统以 NFS 或者 CIFS 方式为客户机提供文件共享服务。采用 NAS 存储虚拟化方案时，每台服务器都是 NAS 存储系统的客户机，它们共享其存储空间和文件。为了提高存储访问的速度，通常为存储访问建立单独的网络，否则将大量占用工作服务网络的带宽。其部署如图 3-2 所示。

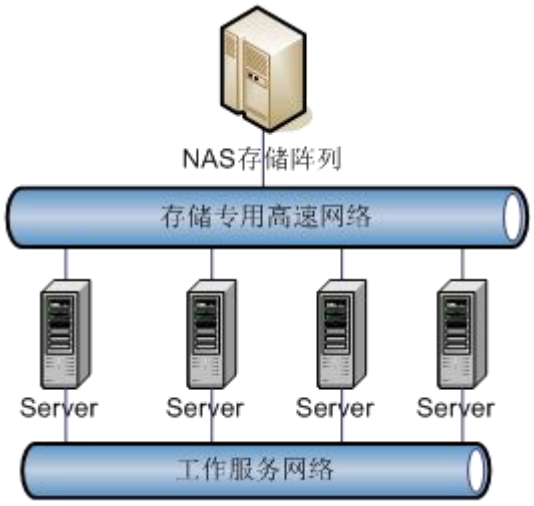


图 3- 2：NAS 存储虚拟化的部署示例

这时，可以在数字有机体虚拟机系统的管理系统中建立网络文件系统池。其中的服务器地址和端口就是 NAS 存储系统的地址和端口。

和 SAN 相比，采用 NAS 存储池的优点是存储空间完全动态分配，缺点是 NAS 系统的性能和扩展性都低于 SAN 系统。同样，它也存在可靠性问题。

3.2.4 数字有机体存储虚拟化

当需要数据共享，系统需要分散部署，存储容量需求大时，就需要采用数字有机体存储虚拟化方案。数字有机体存储虚拟化的方案如图 3-3 所示。

这时，存储设备将直接连接到服务器上，建议采用服务器内置磁盘阵列或者外置磁盘阵列。为了不影响服务网络的性能，建议将内部通信独立在一个网络中。服务器可以分散部署在多个地方，以便实现异地复制，增强系统的抗毁容灾能力。

采用这种方案时，需要使用数字有机体文件系统来管理所有服务器上的存储设备，并形成共享的虚拟文件系统。在这个虚拟文件系统中，再建立数字有机体目录池。系统中的任意

服务器都可以共享这些目录池，从而使得虚拟机可以在系统中迁移，不受存储设备的约束。

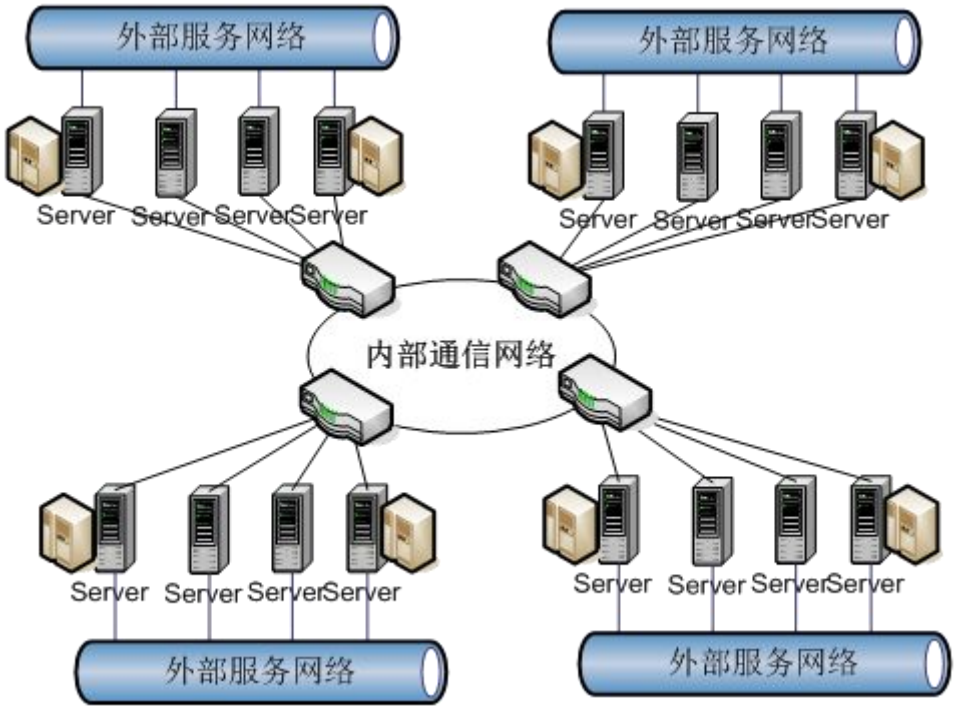


图 3- 3：数字有机体存储虚拟化

3.3 网络虚拟化方案和部署

网络虚拟化需要满足许多方面的需求，如虚拟机通信需求、通信管理需求、通信安全需求等。其次，网络部署方式复杂，涉及的技术和设备众多，不同的需求可能采用完全不同的方案；即使同样的需求，也可以采用不同的技术和方案实现。因此，本文难以全面的说明网络虚拟化的各种技术。下面仅给出一些部署示例。涉及的术语请参照本文 1.4 节的解释。

3.3.1 虚拟机通信方式选择

在考虑网络虚拟化方案前，请先收集以下信息：

1）需要相互通信的虚拟机仅部署在某台服务器内还是跨服务器部署；

2）虚拟机需要和外部网络通信吗？如果需要通信，采用哪种方式：NAT、路由和桥接？

3）需要 VLAN 支持吗？

数字有机体可以为虚拟机提供 6 种网络通信方式，即隔绝网络、NAT、路由、Linux 桥接、OVS 桥接和 Macvtap 桥接。后三种通信方式都采用桥接模式，但是其功能上仍然有差异。通常，将需要相互通信的，采用相同通信方式的虚拟机部署在一个虚拟网络内，本文简称网络。同一个网络内的虚拟机可以相互通信。

1）隔绝网络：即不需要和外界通信的网络。网络内的虚拟机可以相互通信，但是不能和外部网络通信。

2）NAT 网络：该网络通过宿主机的物理接口以网络地址转换（NAT）的方式和外部网络互联。通常，网络内的虚拟机使用私有网络地址，如 192 网段、10 网段的地址。网络内

的虚拟机可以相互通信，和外界通信时则采用 NAT 方式，由网络的网关（实际上是宿主机）完成 NAT 处理，因此可配置 NAT 规则。

3）路由网络(route)：该网络通过宿主机的物理接口，以路由的方式和外部网络互联。实现时，将在宿主机的桥接器上绑定网络的网关地址，由宿主机完成路由。大多数时候，还需要在外部路由器上配置相应路由信息。

4）桥接网络：这种网络通过宿主机上的桥接器与外部网络通信。桥接器使用一个或者多个宿主机的物理接口。根据桥接器的不同，该类网络又可以分为三种：

（1）Linux 桥接网络：使用 Linux 桥接器，是最基本的桥接方式，性能一般。

（2）ovs 桥接网络：使用 Open vSwitch 桥接器，性能更好，且支持 VLAN。

（3）macvtap 桥接网络：使用 macvtap 驱动程序，利用一个以上物理接口作为桥接器。

除桥接网络外，其他的网络都只能局限于单台宿主机内，无法跨宿主机互联。因此当需要跨宿主机部署网络的虚拟机，且虚拟机间需要相互通信时，则必须采用桥接网络。

下表是各种网络的比较。其中，物理接口指是否使用物理接口和外部网络通信。内置服务指是否为子网提供 DHCP、DNS 和网络启动服务，“有”表示无论是否配置都有，“子网”表示需要在子网配置中设定需要启动网关。QOS 指能否限定流量。MAC 指是否可以为桥接器配置 MAC 地址，“可”表示可以配置，“无”表示没有桥接器不能配置，“否”表示不能配置。IP 指桥接器的 IP 地址，即网络的网关地址。

类型	物理接口	和外网互联	内置服务	QoS	MAC	IP	跨主机组网
隔绝	无	无	有	有	可	否	否
NAT	用 1 个	NAT 方式	有	有	可	可	否
路由	用 1 个	路由方式	有	有	可	可	否
Linux 桥接	用 1 个	Linux 桥接	子网	无	否	可	可
ovs 桥接	用 1 个	ovs 桥接	子网	无	否	可	可
macvtap 桥接	用多个	设备桥接	无	无	无	无	可

当上述类型的网络需要同时部署在一台宿主机上，存在着以下限制：

1）隔绝、NAT 和路由类型的网络在同一台宿主机上只能有一个。这里指任意三种类型中的一个。

2）桥接网络和隔绝、NAT、路由网络不冲突，但是不能共享物理网络接口。这要求管理人员仔细规划宿主机上物理接口的使用，避免冲突。同时，也建议宿主机配置多个物理网络接口，以便组网。

通过前面分析可知，隔绝网络，NAT 网络和路由网络只适合单服务器部署方式，在多服务部署时，不要采用它们，除非有特殊的需求。

要注意的是：数字有机体虚拟机系统虽然也支持单服务器部署，但是这时将无法发挥数字有机体虚拟机系统的特性，例如虚拟机迁移、虚拟机抗毁容灾、负载均衡、高伸缩和虚拟机集群等。

要说明的是：下面的例子中虽然只画了两台服务器，但并不表示数字有机体虚拟机系统

建议双机部署，反而是建议至少部署三台以上服务器，甚至支持成百上千台服务器的分散部署模式。

3.3.2 简化的内部私有云

数字有机体虚拟机系统支持各种形式的组网。最简单的方式就是 Flat/DHCP 模式。Flat 的意思是所有虚拟机的网络接口处于同一个平面网中，不存在层次结构。DHCP 的含义是使用动态地址分配方式。当然，平面网中也可以采用固定地址分配方式。

图 3-4 是一种简单的部署结构。这种结构适合于像虚拟实验室这样的无安全隔离需求的内部应用系统。

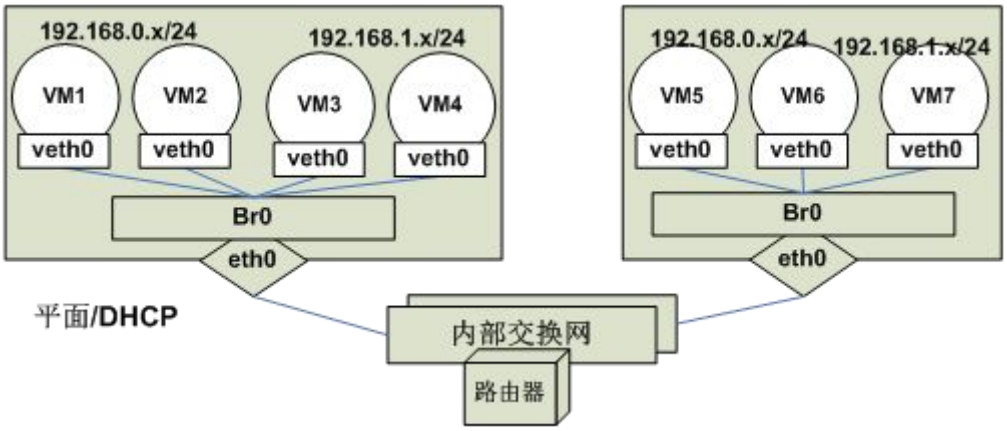


图 3- 4：简单的部署结构

此时，所有的虚拟机都通过桥接器连入交换网中。网络的桥接器都不配置 IP 地址，由物理路由器担任网络网关。如果需要 DHCP、DNS 或者网络启动服务，也都由物理路由器提供。

虚拟机可以处于不同的 IP 子网，并且通过物理路由器相互连通。虚拟机的 IP 地址既可采用 DHCP 动态分配方式，也可以固定指定。唯一的网络隔离措施是在路由器中配置一些防火墙规则以限制 IP 子网间的访问。虚拟机的用户也处于内部交换网中，因此无需连接外部网络的措施。

显然，这种部署结构没有区分用户网络和内部网络，因此网络带宽是共享的，网络的安全性也不足。因此只适合规模较小的，内部使用的场景。

3.3.3 有内外网的示例

如果虚拟机既需要和其他虚拟机通信（内部通信），又需要和外部网络通信（如对外提供服务），则需要考虑如何规划网络。由于虚拟机需要跨宿主机通信，因此必须采用桥接网络来互联（支持 Linux 桥接、OVS 桥接和 Macvtap 桥接）。 这时，所有虚拟机处于同一个平面，相互之间只有 IP 子网隔离，没有 VLAN 隔离。虚拟机通过宿主机的桥接器直接桥接到物理网络，从而可以跨宿主机组网。

如果虚拟机的内部通信网络也需要划分为多个子网时，就需要考虑这些子网间的互联。同样，出于安全考虑，内网和外部网络也常常需要分离，它们之间的通信也需要路由器、防火墙等。网络间的互联既可采用物理路由器实现，也可以由系统的服务器担任网关（即充当路由器）。采用物理路由器的可靠性更好，性能也更优，但是成本更高。采用系统服务器担

任网关将占用服务器资源，不过部署成本更低。

在大多数需要对外服务的网络应用中，都希望将内部通信网络和外部服务网络分离，以便增强安全性，避免带宽争用。图 3-5 是一个服务网络和内部网络分开，由物理路由器实现互联的例子。这需要服务器至少有两个以上的物理网络接口。图中的 eth0 作为外网互联接口，eth1 作为内网互联接口。在系统中分别建立内部交换网的虚拟网络和外部交换网的虚拟网络。虚拟机也部署两个虚拟网络接口，分别使用这两个虚拟网络。内部网络和外部网络间不直接通信。

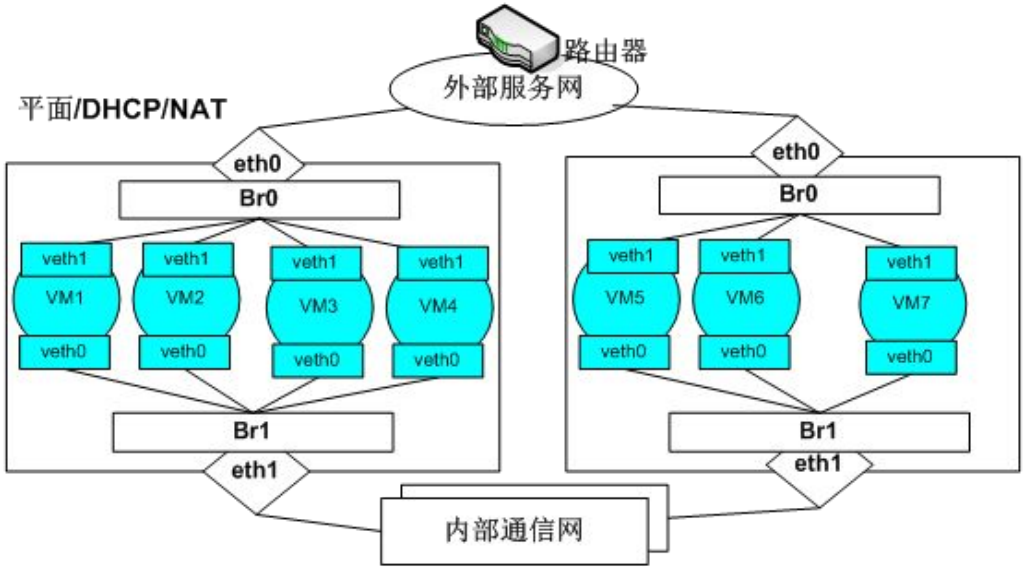


图 3- 5：内外网分离的全物理互联方案

这种方式分离了内外通信，有利于提升虚拟机的通信性能，也可以增强内部通信的安全性。同时，由硬件路由器完成路由或者 NAT，性能更好，减轻了服务器的负载。外部通信网和互联网间还可以根据需要增加安全设备，如防火墙。

如果服务器只有一个物理网络接口，则内外网将无法分离，对外服务的通信将占用内网通信带宽。其结构如图 3-6 所示。这时只需建立一个虚拟网络，供虚拟机间通信和对外通信使用。

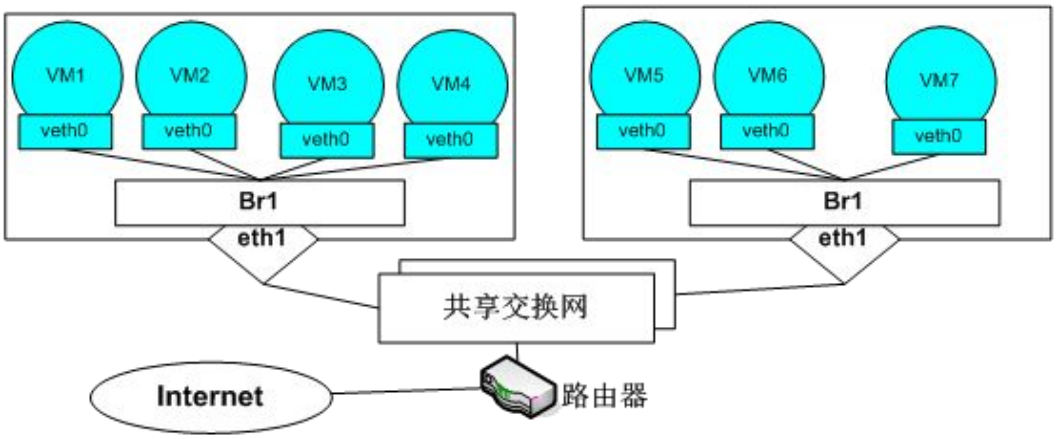


图 3- 6：内外网不分离的全物理互联方案

下面说明使用系统服务器担任网关的部署方式。数字有机体虚拟机系统支持自动部署子网网关，即由系统选择（也可指定）一台服务器担任子网的网关。子网和外部服务网络的互

联通过网关完成。

内部子网和外网的通信方式有两种，即 NAT 方式和直接路由方式。两种方式的部署结构是相同的，不同的仅仅是担任网关的宿主机如何处理外部交换网和互联网间的通信。

如果服务器网络接口充足，如有三个物理网络接口，则可以将外部通信网络和内部通信网络隔离，使用一个物理网络接口做内部网络桥接器，另一个网络接口则作为和 Internet 互联的桥接接口。部署结构如图 3-7 所示。虚拟机的虚拟网卡都通过宿主机的物理网卡桥接到交换网。可为不同业务的虚拟机建立单独的子网（安全组）。每个子网处于不同的 CDIR（子网段）中。

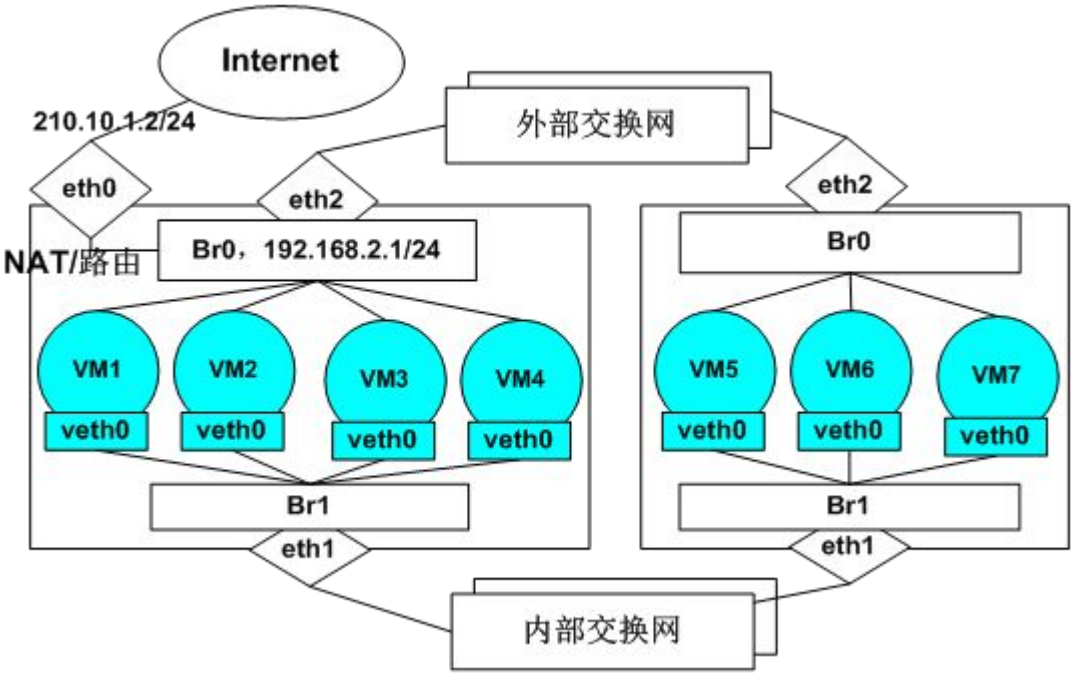


图 3- 7：内外网分离服务器做网关的方案

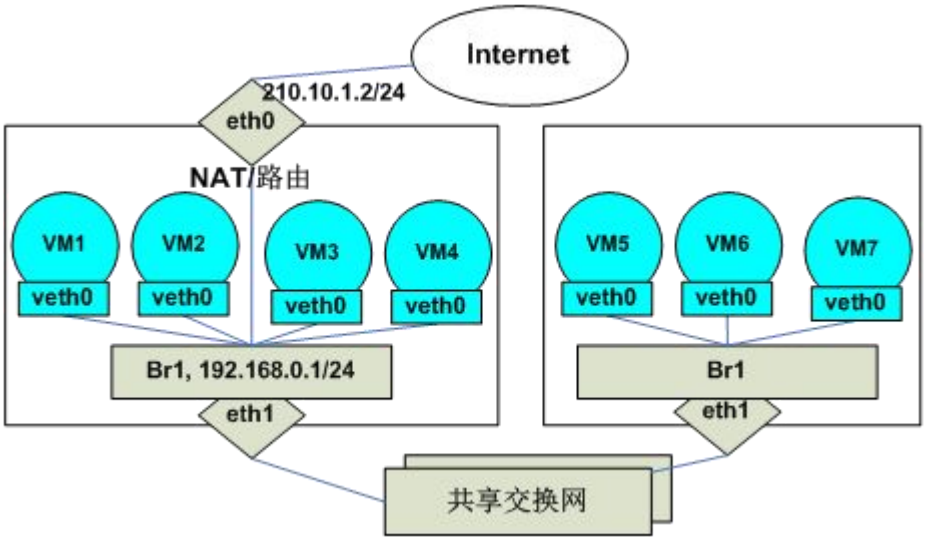


图 3- 8：服务器做网关的 NAT 模式

如果宿主机接口不足，则可以将外部交换网和内部交换网合并，其部署结构如图 3-8

所示。例如图中的场景，虚拟机使用 192.168.0.0/24 网段互联。虚拟机的网卡地址采用 DHCP 方式分配。宿主机上只配置一个桥接设备即 Br1。所有虚拟机的虚拟网络设备（tap）都通过宿主机的 eth1 由 Br1 桥接到共享交换网。这样同一个业务间的虚拟机就可以通信了。

宿主机的 eth0 配置公网地址，扮作 NAT 网关或者路由器，实现虚拟网络和 Internet 的互联。如果是 NAT 网关，系统将在宿主机的 IP-Table 中增加安全组的 NAT 规则。这组规则中包括将业务的外网浮动 IP 映射到虚拟机私有地址的规则。如果服务器扮作路由器，则需要增加路由规则。

如果采用 NAT 网关，互联网用户的请求发送到业务的 VIP。绑定 VIP 的宿主机的内核将收到该请求，然后按照设定的规则，使用 DNAT 将请求转发给内部的虚拟机。内部虚拟机的响应也通过绑定 VIP 的宿主机再转发出去。

3.3.4 VLAN

当在同一个系统中部署多个业务的虚拟机时，可能希望进一步隔离不同业务间的虚拟机通信，以增强安全性。这时，需要采用 VLAN。数字有机体虚拟机系统支持 Open Virtual Switch（OVS）虚拟网络，可以通过它构建支持 VLAN 的虚拟交换网，从而达到隔离不同业务的目的。

采用基本的 VLAN 技术时，单个交换系统最多只能支持 4096 个 VLAN，因此通常认为 VLAN 模式只适合私有云。Flat 没有这个限制，因此被认为适合于公有云系统。不过，Open Virtual Switch 也支持 VxLAN，即扩展的 VLAN，它可以支持更多的 VLAN 子网。

3.4 虚拟机部署

3.4.1 需求分析

在考虑虚拟机部署之前，请先收集虚拟机需要的资源信息，例如以下这些信息：

1）CPU 需求：结构（X86-64、I686、IA32 等），核心数。如果需要，可能还包括具体型号，例如 PC1.0 等。

2）内存需求：容量大小。

3）存储需求：容量、性能要求、可靠性要求和共享需求。

4）网络需求：网络接口数量，每个接口需要连接的物理网络，QOS 需求。

5）监视需求：远程监视还是宿主机上监视，分辨率，是否有流畅性要求，例如播放视频就要求流畅性好。

6）I/O 接口需求：需要模拟的 USB、串口等 I/O 接口。

7）运行可靠性需求：需要迁移吗？需要离线备份吗？需要快照吗？

8）虚拟机协作要求：是否需要和其他虚拟机通信，协同工作？

对 CPU 需求，第一要注意的是：在宿主机上模拟不同于宿主机 CPU 结构的 CPU 需要完全虚拟化，其性能损失是必然的。反之，如果模拟相同结构的 CPU，则可以采用半虚拟化技术，例如 KVM。半虚拟化技术能够提供更好的性能。第二要注意的是：虚拟机运行时

是不能调整其 CPU 结构、类型和核心数等的，只有关闭虚拟机后才能修改它们。第三，CPU 虚拟化并不能模拟不同的 CPU 频率。第四，在迁移虚拟机时要考虑目的和源宿主机的 CPU 差异，在不同 CPU 结构的宿主机间迁移虚拟机可能失败。第五，虽然宿主机可以模拟出远大于实际 CPU 核心数的虚拟 CPU 核，但当模拟的虚拟 CPU 核数超过实际的 CPU 核心数时，虚拟机间将相互争用资源，使得虚拟机的处理能力下降。

对内存需求，需要注意的是：当宿主机的物理内存耗尽时，系统的性能将急剧下降，因此要避免在宿主机上运行过多的虚拟机，以致耗光宿主机的物理内存。换句话说，宿主机上运行的虚拟机的内存总数应小于宿主机的内存容量，且要预留足够的空间给宿主机操作系统使用。建议预留不小于 2GB 的内存容量。

存储需求除了要考虑容量外，也要考虑性能、可靠性和共享需求。由文件系统实现的虚拟磁盘存在性能损失（特殊情况下也能提升性能，如缓存加速、并行读写）。可靠性方面，文件系统缓存将增大丢失数据的风险，但文件系统也可以提供数据复制、异地复制、故障屏蔽等功能，又能提升数据可靠性。共享方面，物理设备方案难以实现数据共享，而文件系统则容易实现数据共享。当需要在虚拟机间共享数据时，文件系统实现的存储虚拟化是必然的选择。

网络需求的分析参见网络虚拟化部分，这里不讨论。

监视需求：当一台宿主机上同时部署多个虚拟机时，难以同时在宿主机上进行监视，因此网络远程监视是必须的选择。在宿主机上监视时，显示的分辨率可以更大，显示的流畅度也很好。网络远程监视时，受限于网络带宽和延迟，难以实现高分辨率的监视，且流畅度也较差，不适宜观看视频等用途。远程监视时可能还存在数据加密传输需求，不过要注意加密带来的大量计算开销。大多数用于网络服务的虚拟机都不需要随时监视着，因此远程网络监视是常见的选择。

I/O 接口需求：在虚拟机中使用 I/O 接口并不容易，原因是 I/O 接口多数难以远程模拟，更多时候是将宿主机上的 I/O 接口共享给虚拟机使用。

运行可靠性需求：如果虚拟机要使用 I/O 接口，或者要使用无法远程访问的网络接口和存储设备，则虚拟机只能在固定的宿主机上运行，无法迁移，也无法复制到其他服务器上运行。反之，则虚拟机可以迁移到不同的服务器上运行。为虚拟机制作快照、迁移虚拟机、保存和恢复虚拟机都可以增强虚拟机的可用性，即在当前运行宿主机故障后，仍然可以在其他宿主机上运行该虚拟机。

虚拟机协作需求：如果虚拟机需要和其他虚拟机通信，则它们之间存在协作关系。需要相互协作的虚拟机最好部署在同一个虚拟网络中，以便相互通信；甚至最好部署在同一台宿主内，以便减少物理网络通信，提升通信速度。当需要协作的虚拟机部署在不同的宿主机上时，虚拟网络就需要跨宿主机部署，因此必须采用桥接网络。

在明确虚拟机的需求后，即可考虑虚拟机的部署方式，并在系统中配置它们。

3.4.2 虚拟机部署

在系统中部署虚拟机的基本思路是：先部署虚拟机运行环境，然后部署虚拟机。为此，需要部署虚拟机需要的以下运行环境：

1）项目：这里的项目既可能是一个虚拟的概念，也可能真实对应现实中的一个项目。

不管怎样，数字有机体虚拟机系统中的虚拟机都必须归属于某个项目。因此，先要配置某个项目，以便在项目中创建虚拟机。当然，一个项目可以有许多个虚拟机。

2）存储池和存储卷：如果虚拟机需要使用存储设备，就应当为其部署一个存储卷。每个存储卷必然是属于某个存储池的，因此需要先新建存储池。不同的存储虚拟化方案需要建立不同的存储池。

3）镜像文件：如果虚拟机需要使用某个文件来模拟 CD-ROM，则先上传该镜像文件到系统。

4）网络：如果虚拟机需要使用虚拟网络，则需要先配置虚拟网络。这可能包括配置虚拟网络使用的子网，以及网络接口使用的安全组（即安全规则）。

5）安全对象：如果存储池或者存储卷需要密钥来访问，例如密码或者证书。就需要在新建存储池或者存储卷前建立安全对象。安全对象用于保存密码和证书等。

在部署好运行虚拟机需要的环境后，即可部署虚拟机。本系统将一个完整的虚拟机配置称作虚拟机实例（简称实例）。用一个虚拟机实例可以启动多个虚拟机，每个虚拟机都可以独立的运行。

用一个虚拟机实例启动多个虚拟机时，需要考虑虚拟机使用的资源的共享问题。同一实例的虚拟机使用的各自独立的虚拟网络接口，只是连接到相同的网络，因此不存在资源使用冲突问题。需要考虑的是虚拟机使用的存储设备，无论是存储卷还是镜像文件。严格的说，如果存储卷用作实例的存储设备，则除非存储卷是只读的，否则都将导致存储卷被破坏。类似的，如果两个虚拟机同时修改镜像文件也会出现这样的情况。不过，可以将镜像文件本身设置为只读的，但在虚拟机实例配置中设备却是可读写的。这时，系统将自动为镜像文件建立一个临时快照，读写操作针对临时快照进行。不过，虚拟机停止运行后所有的修改都将丢失，因为系统自动销毁临时快照。

实际上，一个实例同时启动多个虚拟机的方式主要用于虚拟机集群，其目的是用多个虚拟机并行提供服务。这时，服务需要的数据保存在一个共享的数据库或者文件系统中，而虚拟机使用的存储设备只用于存储程序和临时数据，因此停止虚拟机后丢失临时数据并没有什么问题。

3.5 虚拟机集群

计算资源动态分配存在两个问题。第一个问题是运行中的虚拟机不能动态增减计算资源，即不能增减 CPU 核心数或者修改类型。如果某个业务只使用一个虚拟机，则在业务繁忙时也无法再为其增加计算资源，除非关闭虚拟机重新配置后再启动。第二个问题是单个虚拟机的计算能力不可能超过宿主机的计算能力，当业务需要的计算能力超过单台服务器的能力时，单纯的虚拟机技术无法解决这个问题。

针对上述两个问题，数字有机体虚拟机系统提供虚拟机集群支持。简单地说，就是利用集群技术，将多个虚拟机整合起来提供共同服务。单个虚拟机服务能力不足时，系统再启动一个虚拟机，和原来的虚拟机组成集群，如果原来的虚拟机集群能力也不足，则可能在其他服务器上启动虚拟机，以扩展服务能力。这样就解决了上述两个问题。

虚拟机集群的部署方式有两种，一种是利用负载均衡器部署的虚拟机集群。第二种是用数字有机体虚拟机系统自身功能部署的虚拟机集群。下面分别予以说明。

3.5.1 独立负载均衡器的虚拟机集群部署方案

如果已经有网络负载均衡器，则可以利用已经部署的网络负载均衡器来部署虚拟机集群。这时，对负载均衡器来说，虚拟机集群中的虚拟机和普通物理服务器没有差别。不同的是，负载均衡器只能使用标准的 IP 通信来监控服务节点的成活状态，幸好这是网络负载均衡器都支持的。

当采用单独的负载均衡设备构建虚拟机集群时，部署示意图如图 3-9 所示。物理的负载均衡设备——负载均衡器部署在外部交换网络和互联网间，通过网络负载均衡功能向虚拟机分配负载。这时，系统中的虚拟机对负载均衡器来说就是真实的服务节点，由它完成故障检测和请求调度功能。

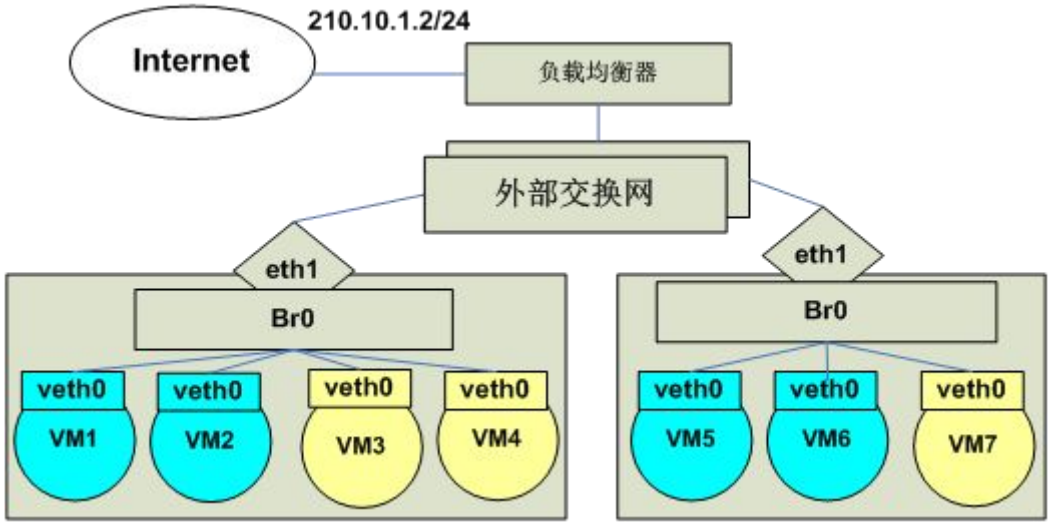


图 3- 9：部署物理负载均衡设备的虚拟机集群方案

采用这种方式时，数字有机体虚拟机系统仅仅负责检测集群中虚拟机的负载情况，以决定是否增加或者减少虚拟机的数量。其余的工作由负载均衡设备完成。当现有的虚拟机都很忙时，系统可以自动增加虚拟机，从而分担负载。这些虚拟机可以运行在不同的宿主机上，从而实现服务能力扩展。

由于虚拟机需要运行在不同宿主机上，因此需要采用桥接方式来为虚拟机提供通信接口。虚拟机的外部通信接口通过宿主机上的桥接器直接接入外部交换网络(可以认为是链路层连通的交换网)。

不同的负载均衡器有不同的组网要求，其工作原理也各不相同，因此具体的部署方案需要根据具体的负载均衡器来设定。本文不深入讨论其细节。

3.5.2 数字有机体虚拟机集群部署方案

第二种方案是由数字有机体系统来完成负载均衡器的功能。这时，系统中的某台服务器将扮演负载均衡器的角色。和独立负载均衡器不同的是，当前扮演负载均衡器的宿主机故障时，系统可以自动或者手动选择其他宿主机继续担任负载均衡器的角色，从而提升系统的可靠性。当系统的服务器数量充足时，系统的可靠性远远高于独立负载均衡器的方案。

由于由数字有机体服务器担任负载均衡器的角色，因此负载均衡的配置参数也需要设置，这包括服务的虚拟 IP 信息和监控虚拟机存活的方式。每个服务都使用单一的公网 IP 对外提供服务，其配置信息即虚拟机集群中的 VIP 参数。数字有机体虚拟机系统支持三种监控虚拟机存活的方式，即 ping、tcp 连接和 http 请求。可以根据需要选择合适的监控方式。

通常，用户的网络请求先到达服务子网的网关，然后到达担任负载均衡器角色的宿主机（以下简称负载均衡节点），最后再到达具体处理请求的虚拟机。从网关到负载均衡节点的处理是标准的 IP 路由，即网关先通过 ARP 请求解析服务 IP 的 MAC 地址，然后将 IP 包转发到相应的 MAC 地址。因此，服务 IP 子网的网关具体在哪里并不重要，它可以在网络运营商那里，也可以在数字中心的互联网络中，甚至可以就由负载均衡节点做网关。下面的部署示意图中都不在单独画出服务 IP 子网的网关。

数字有机体虚拟机系统支持两种网络请求转发方式，即网络地址转换（NAT）和直接路由。采用不同的转发方式需要的部署方案也不同。下面分别讨论它们。

3.5.2.1 基于 NAT 的数字有机体虚拟机集群

这种方案只需要一个公网 IP 地址，所有的虚拟机都使用私有地址提供服务。它适应了当前 IP 资源紧缺，网络服务安全要求高的应用环境，因此被广泛采用。

图 3-10 是这种方案的部署示意图。这时，用户请求的地址，即服务 IP 地址，绑定在负载均衡节点的互联网网络接口上，用于接收用户请求。为了让每台宿主机都可担当负载均衡节点，在互联网接入链路上部署一台交换机，让所有宿主机的 eth0 都连接到该交换机。同一个时刻，系统中只有一台宿主机担任这个虚拟机集群的负载均衡节点，因此公网 IP 地址只会绑定在当前担任负载均衡节点的宿主机的 eth0 上。

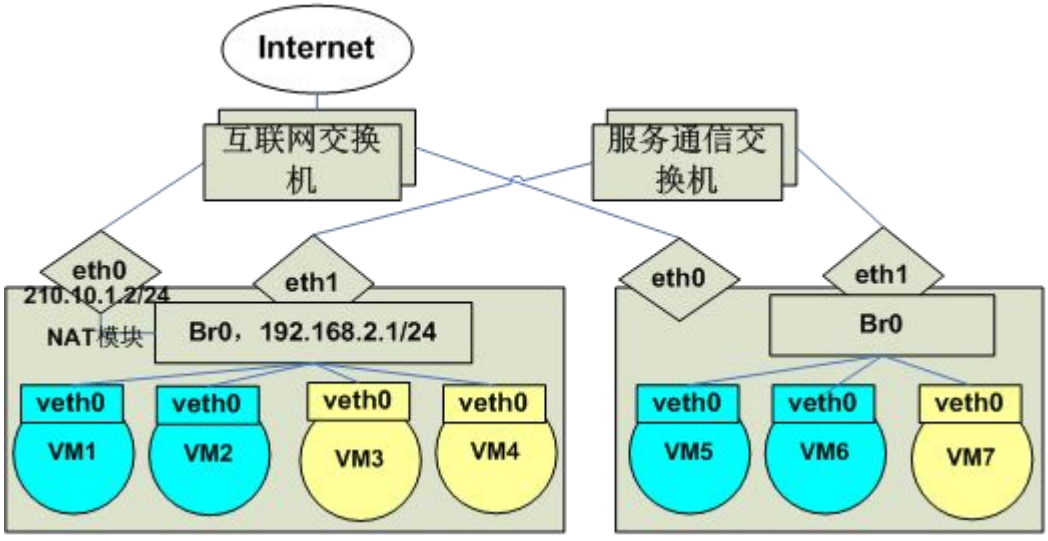


图 3- 10：数字有机体服务器构建的 NAT 虚拟机集群

虚拟机通过桥接器，由 eth1 连接到服务通信交换机，形成一个内部通信子网。虚拟机使用私有地址进行服务，如图中的 192.168.2.0/24 网段地址。虚拟网络的网关需要和负载均衡节点在同一个宿主机上，以便完成地址转换，因此虚拟网络的网关地址绑定在负载均衡节点的虚拟网络桥接器上。

用户的请求先到达负载均衡节点，即图中的 eth0 接口。内核 IPVS 模块先确定请求应该转发的节点，建立转发记录，然后改写请求消息中 IP 头的源地址为 eth0 绑定的地址，最后将请求转发给虚拟机的网络接口；虚拟机处理请求后将响应发给负载均衡节点，由负载均衡节点转发给用户。

数字有机体虚拟机系统可以启动 DHCP 服务和 DNS 服务来满足子网的配置需求。因此，虚拟机的 IP 地址可以通过 DHCP 方式获得。这样自动启动虚拟机时可以自动分配地址，无需人工干预。而且虚拟机内无需再配置其他地址，因此完全可以自动启动虚拟机提供服务。

当负载均衡节点故障时，可以在其他宿主机上启动新的负载均衡服务，从而使系统继续提供服务，这提升了系统的可靠性。

3.5.2.2 基于直接路由的数字有机体虚拟机集群

采用直接路由(Direct-Route)方式时，返回用户的响应无需通过负载均衡节点，因此可以减少负载均衡节点的开销，从而使得集群的规模可以更大。但是，它要求每个虚拟机都有一个和服务 IP 地址同网段的 IP 地址，以便响应可以直接返回给用户。

采用直接路由的部署结构如图 3-11 所示。需要为服务 IP 地址所在的子网建立一个虚拟网络，虚拟机使用该网络和互联网通信。建议部署独立的路由器作为服务子网的网关，并在路由器上配置 DHCP 服务，以便虚拟机的网络接口可以自动获得 IP 地址。

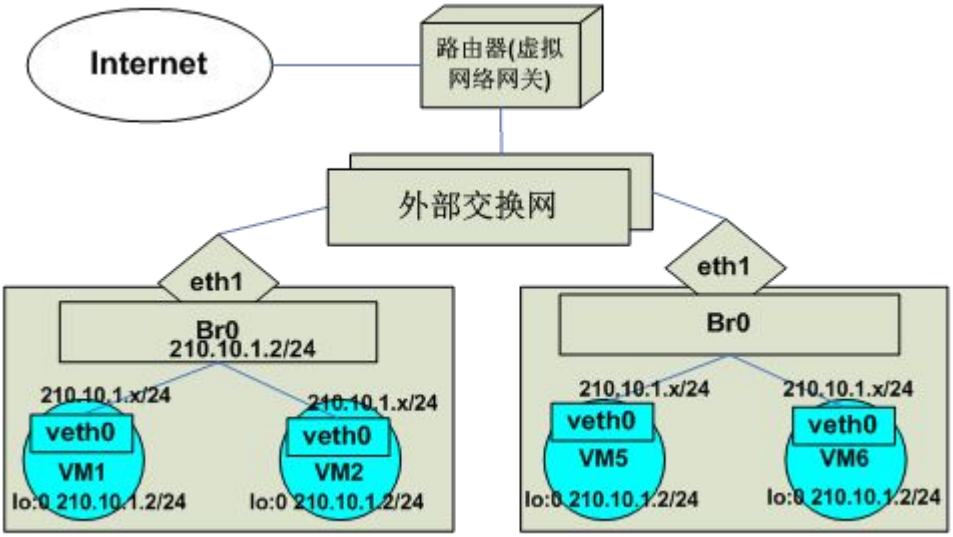


图 3- 11：数字有机体服务器构建的直接路由虚拟机集群

用户请求的服务 IP 地址绑定在负载均衡节点的虚拟网络的桥接器上。当负载均衡节点转移为其他宿主机时，该 IP 地址也相应的转移绑定到新宿主机的桥接器上。只有负载均衡节点的服务 IP 地址是对外响应 ARP 请求的，以便用户请求只被转发到负载均衡节点。

和 NAT 方式不同的是，每个虚拟机内也要绑定服务 IP 地址，不过该地址要求绑定在回环设备 lo 上，作为 lo 的别名。例如 ifconfig lo:0 服务 IP 地址/32。注意，绑定 IP 地址时的掩码为 255.255.255.255。默认回环设备上的地址不会用于响应 ARP 请求，因此绑定的 IP 地址仅仅用于虚拟机接受负载均衡节点转发来的用户请求。

虚拟网络需要使用和服务 IP 地址同网段的 IP 地址，即每个虚拟机都要有一个公网地址

（下面称其为通信地址），且和服务 IP 地址同网段。通信地址绑定在虚拟机的 eth0 上，即连接虚拟网络的接口上。

用户请求通过网关到达负载均衡节点（因为服务 IP 地址绑定在负载均衡节点的桥接器上，由他对外响应 ARP 请求），在经过调度后直接路由到虚拟机网络接口。虚拟机因 lo 上有服务地址，因此接受请求并进行处理。虚拟机的响应消息用服务地址作为源 IP 地址，直接通过路由器返回给客户。

如果不部署物理路由器，也可以启动子网的自动网关部署功能，由数字有机体服务器扮演虚拟网络的网关角色。这时，虚拟网关地址也绑定在桥接器上。

3.6 多站部署示例

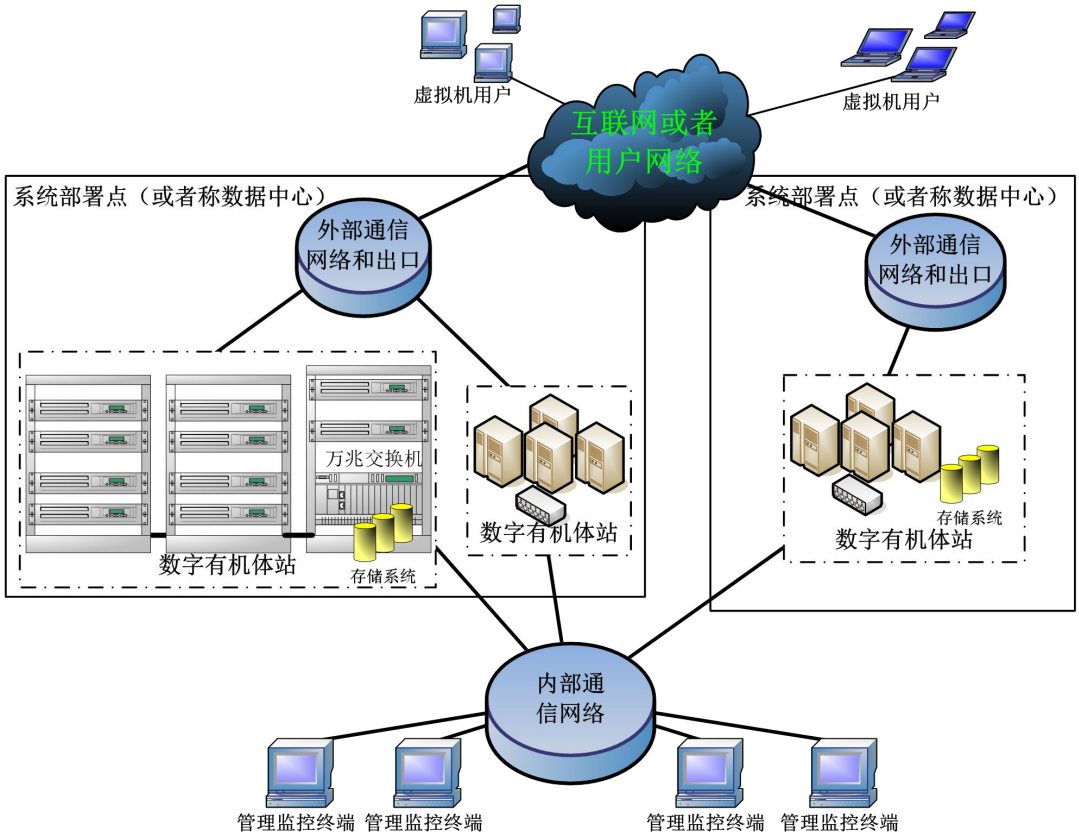


图 3- 12 数字有机体云部署示意图

数字有机体虚拟机系统的典型部署场景如图 3-12 所示。系统部署在多个部署点。每个部署点的规模可大可小。规模大的就是一个数据中心，规模小的可能就是几台或者几十台服务器。这些部署点间通过高速网络互联，如 10G 光纤网络。在部署点内部，根据规模不同，需要建立相应的内部传输网络。

考虑到系统内服务器间的通信量大，为了避免和服务争用网络带宽，最好分别建立外部网络和内部网络。而且，系统中将运行大量的虚拟机，这些虚拟机都需要和外部网络通信，需要的带宽较大。同时，这些虚拟机需要频繁访问内部资源，需要的带宽也很大。这样，每

台服务器（即运行虚拟机的宿主机）都需要安装多块网卡，分别用于与内外网络的通信。

系统中服务器的性能、大小都存在差异。可以认为是由新旧、大小、品牌各不相同的服务器构成。这些服务器处于网络中的不同网段。同一网段的服务器构成一个数字有机体站。不同网段的服务器无法组成一个数字有机体站。这些服务器可以集中部署在同一个或者多个机柜中。如果希望准确的感知网络位置，可以将同一个交换机下的服务器组成一个站。在一个数据中心中可以部署许多数字有机体站。

每台服务器都内置有磁盘或者外接存储系统。各台服务器的存储容量、性能各不相同。每台服务器的存储空间可以分为三个部分：

1）宿主机系统空间。这个空间用于安装宿主机的操作系统，即安装数字有机体系统程序。

2）虚拟机镜像临时存储空间：用于存储正在运行的虚拟机的临时系统镜像文件快照。虚拟机不再运行时，相应的快照也被删除，从而腾出空间给其他虚拟机使用。

3）共享存储空间：共享到数字有机体虚拟存储系统的空间。

这三个空间要求是相互独立的。他们可以是同一设备上的不同分区，也可以是不同设备甚至逻辑卷上的分区。

所有服务器的共享存储空间将被数字有机体文件系统使用，形成虚拟统一的逻辑存储空间，供所有虚拟机和宿主机使用。

虚拟机系统的用户可以初步分为两类，即虚拟机用户和管理用户。虚拟机用户在外部网络上使用和访问虚拟机，操作所谓的按需分配的虚拟主机。每个系统部署点（数据中心）都有自己的外部网络出口。这些网络出口的带宽可能大小不同，所处的网络位置也可能不同。典型的，一个数据中心可能同时有电信、联通、移动等的网络出口，而某个小的部署点可能只有一个电信出口。而且这些部署点因地理位置不同，即使是接入同一个网络运营商（如电信）的网络出口，对虚拟机用户来说，其带宽也是不一样的，因此其使用体验也是不同的，尤其对那些需要较高带宽需求的虚拟桌面应用。

管理用户则只能从内部网络访问系统，因此他们都只连入内部通信网络。管理用户可以通过多种手段访问系统，如浏览器、SSH、FTP，甚至远程桌面等。系统管理用户又可分分为两类，即系统管理员和项目管理员。

4. 软件安装和启动

当前，数字有机体虚拟机系统的软件完全集成在数字有机体系统中。您可以通过安装4.3 以上版本的数字有机体系统而获得数字有机体虚拟机系统软件。

4.1 系统运行环境要求

系统安装的最低配置要求如表 4-1 所示。

表 4-1 硬件最低配置

处理器	Intel 至强 1.6G 4 核心或以上处理器，也支持同级别的 AMD 处理器
内存	4G 或以上内存
硬盘	40G 或以上硬盘
网卡	100/1000M 网卡

不过，要很好的运行数字有机体虚拟机系统，根据当前硬件设备的发展水平，建议的服务器配置如 4-2 所示。

表 4-2 推荐的服务器配置

整机	中级服务器系统
处理器	双路或者 4 路，Intel 至强 2.0G 6 核心或以上处理器，也支持同级别的 AMD 处理器
内存	256G 或以上服务器内存
硬盘	系统盘建议配置为 600GB 的 SAS 磁盘。内置磁盘阵列卡或者外接磁盘阵列，支持 raid0、raid5。存储容量根据需求配置。
网卡	4 个或以上 1000M 或者万兆网络接口
网络	千兆以上交换网络，最好是万兆交换网络

具体的部署方案请参加本文第三章的描述。

4.2 运行环境准备

在部署系统之前，请参照本文第三章的内容，仔细规划系统的部署方案。建议在完成以下准备工作后再安装数字有机体虚拟机系统。

规划部署方案：在部署系统之前，建议仔细规划。这包括规划如何部署网络，包括网络结构、子网划分、地址分配、要提供哪些服务，网络带宽分配等；并记下这些规划信息，以便在配置虚拟网络时可以获得。也包括规划虚拟存储系统，如存储设备部署方式、访问方式、空间分配等。

准备系统网络环境： 首先按照系统规划，部署系统运行的网络环境。这可能包括部署网络路由器、交换机、防火墙等设备，也包括网络布线，连通要运行数字有机体虚拟机系统的各台服务器。

准备存储环境：如果要使用网络存储设备，例如 IP-SAN 或者 NAS 存储系统，请先完成这些存储设备的部署，并在这些系统中配置给数字有机体虚拟机系统使用的存储单元，以便数字有机体虚拟机系统可以访问它们。注意记下访问这些存储设备的参数，如网络地址、端口、访问用户名和密码等。如果只使用服务器内置的磁盘系统或者外接的磁盘阵列，也请先完成这些设备的配置。例如配置服务器上内置磁盘的阵列，或者配置外接磁盘阵列。

收集服务器信息：在安装软件前，请收集服务器的以下信息：

1）处理器信息：CPU 的结构、CPU 数量、每个 CPU 的核心数，具体的 CPU 信号。

2）内存信息：每台服务器的内存容量。

3）网络接口：确定服务器每个网络接口的编号，每个网络接口的用途和相应的网络连接是否正常，收集每个网络接口的带宽信息、地址分配、使用方式等。

4）存储信息：确定要安装系统软件的磁盘，以及磁盘空间的分配方案。建议使用独立的磁盘安装系统软件，以免应用系统意外覆盖系统磁盘。收集各个磁盘或者逻辑设备的容量、性能和使用方式等信息。

5）服务器部署规划：数字有机体虚拟机系统支持大量服务器的分布式部署。在同一个子网部署的，相互紧密协作的服务器被组织为一个站，两个站可以在不同的子网，中间通过路由器等网络连接。系统支持同时部署许多个站，并让这些站协同工作。

4.3 软件安装步骤

需要在每一台服务器上安装数字有机体虚拟机系统软件。不要期望安装完一台服务器后，其他服务器就可自动安装上系统软件并完成配置。不过，在需要大量部署时，仍然有一些提升安装速度的方法。例如，如果每台服务器都使用独立的系统磁盘，且这些服务器的硬件配置都是相同的，则可以在安装完一台服务器后，其他服务器的系统磁盘都由安装好的系统盘克隆出（使用 ghost）。另外，数字有机体虚拟机系统也支持网络安装。在完成一台服务器安装后，可以在该服务器上配置网络安装服务，其他服务器即可通过网络并行的安装。当使用千兆网络时，其安装速度要高于使用光盘安装。无论如何安装，每台服务器仍然需要手动进行配置。这包括配置服务器的网络接口、数字有机体系统参数和虚拟机守护进程参数等。

数字有机体虚拟机系统的安装总体上可以分为三步。第一步是安装数字有机体系统（含数字有机体工作平台、数字有机体工作库、数字有机体虚拟机系统软件，以及其他集成的模块）。第二步是配置数字有机体系统。第三步是配置数字有机体虚拟机系统。下面分别进行说明。安装数字有机体系统和配置数字有机体系统请参考 4.3 以上版本的《数字有机体系统安装指南》，这里不再详细描述。

如果尚未安装数字有机体虚拟机系统，您必须首先获取它的安装包 dosvm.war 和 dosvmd_1-2_amd64.deb；然后再安装它们，安装方式如下：

1)将 dosvm.war 网站安装包拷贝到 tomcat 服务器的应用目录/usr/local/tomcat/webapps/；

2)执行“dpkg　-i　dosvmd_1-2_amd64.deb”命令安装 dosvmd 程序。

4.4 软件配置

安装到数字有机体系统中的虚拟机系统软件主要分为两个部分。一个是每台服务器上都必须运行的守护程序，即 dosvmd。可以使用命令“dpkg –l dosvmd”查看软件包的信息。安装后的执行文件为/usr/local/bin/dosvm/dosvmd，/usr/sbin/dosvmd 是它的符号链接。

另一部分软件是虚拟机系统的管理系统。它是一个在 tomcat 中运行的网站。软件包名称为 dosvm.war。安装后，位于/usr/local/tomcat/webapps/目录。启动 tomcat 后将自动解开它，并放在/usr/local/tomcat/webapps/dosvm/目录下。

数字有机体虚拟机系统支持通过 SSL 方式加密通信消息。使用 SSL 通信需要为通信的双方签发安全证书。数字有机体虚拟机系统的通信主体包括：每台服务器上的守护程序 dosvmd、管理网站的 tomcat 服务程序和管理人员使用的浏览器。因此，如果要使用加密安全通信，就需要为每台服务器的守护程序制作和配置安全证书，并为网站的 tomcat 服务程序制作和配置证书。此外，libvirtd 程序也需要使用 SSl 证书。

注意：https 不是强制使用的，不启用则没必要制作 https 的证书。dosvmd 程序和 dosvm 管理网站在正常工作时需要通信，因此在制作安全证书时，必须是同一个根证书，并且是否使用 SSL 的策略也必须保持一致，要么两者都使用 SSL 安全证书，要么两者都不使用。

4.4.1 建立虚拟机系统数据库

数字有机体虚拟机系统需要使用一个共享的数据库，用于保存用户的配置信息。该数据库需要建立在数字有机体工作库系统（即数据库管理系统）中。以便所有服务器都能访问它。

您只需在一台服务器上建立虚拟机系统的数据库即可。所有的服务器都将通过数字有机体工作库系统共享该数据库。

在要建立虚拟机系统数据库之前，请先启动数字有机体工作库系统。你可以通过 dossql 程序登录数字有机体工作库系统，并确认它已经正常运行。

安装好的数字有机体系统中，在/usr/local/bin/dosvm/目录下，有数据库定义文件 dosvm.sql。可以直接通过 dossql 程序执行该定义文件来创建虚拟机系统的数据库，也可以通过该目录下的 install-db 脚本程序来创建虚拟机系统的数据库。在服务器的命令终端中，可以使用如下命令来启动创建数据库的脚本：

```
cd /usr/local/bin/dosvm/db/
./ install-db
```

脚本运行的第一个对话框(图 4-1)用于输入要创建的虚拟机系统数据库的名称（注意：目前只能使用 dosvm）。请记下这里配置的数据库名称、用户名和口令等信息，在配置守护程序 dosvmd 的参数时将使用它。

图 4- 1：输入虚拟机系统数据库名称

确认后将提示输入访问数据库的用户名(图 4-2)。你可以为该数据库新建一个用户，也可以使用已有的用户。该脚本将尝试新建用户并为其授权。

图 4- 2 输入虚拟机系统数据库的用户名

确认用户名后，将提示输入访问数据库的密码（图 4-3）。

图 4- 3 输入虚拟机系统数据库访问密码

确认访问密码后，脚本将做最后的确认(图 4-4)，这时还可以放弃数据库的创建。如果选择“是”则将创建数据库并授权。

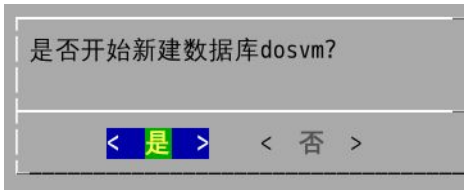


图 4- 4 确认数据库创建

如果数据库创建成功，将给出如下的提示(图 4-5)，否则将显示错误信息。

图 4- 5 创建数据库成功的提示信息

可以使用 dossql 程序访问数字有机体工作库系统，查看创建的数据库的状态，数据库位置和访问数据等。

4.4.2 配置虚拟机系统管理网站的数据库

虚拟机系统管理网站由服务器中的 tomcat 运行。初始发布包为 dosvm.war，放在 tomcat 的 web 应用目录下，即/usr/local/tomcat/webapps 目录。启动 tomcat 服务时，tomcat 将自动解压该文件。

运行虚拟机管理网站只需配置其中的虚拟机数据库访问参数。配置文件为/usr/local/tomcat/webapps/dosvm/WEB-INF/applicationContext.xml。需要使用编辑器修改其中的<bean id="dataSource">节中的 property name="url"的值。其格式如下：

<property name="url"

value="jdbc:mysql://localhost:3306/dosvm?useUnicode=true&characterEncoding=utf-8">

</property>

其中的 localhost 是数据库服务器的访问地址。如果服务器在数字有机体工作库系统中，可以用 localhost 作为访问地址，否则请填写数字有机体工作库系统中的某台服务器的 IP 地址。“3306”是访问的端口号。通常数字有机体工作库使用 3306 作为服务端口号。除非设定过服务端口，否则无需修改它。其他的参数请不要修改。

完成配置后，需要重新启动 tomcat 服务，以便新的设置起效。重启 tomcat 服务可以使用如下的命令：

service tomcat7 restart

注意服务器上 tomcat 的服务名称。可以在/etc/init.d 目录下看到 tomcat 服务的启动脚本。如果脚本文件的名称不是 tomcat7，则使用脚本文件的名称作为服务名。

4.4.3 配置 libvirtd 程序

Libvirtd 程序的配置文件有三个，分别是“/etc/default/libvirtd”、“/etc/libvirt/libvirt.conf”和“/etc/libvirt/libvirtd.conf”。如果修改了配置文件，重启有效，重启的命令是“service libvirtd restart”。

配置文件“/etc/default/libvirtd”需要修改参数“start_libvirtd”和“libvirtd_opts”，修改后应该为：

```
# Start libvirtd to handle qemu/kvm:
start_libvirtd="yes"
# options passed to libvirtd, add "-l" to listen on tcp
libvirtd_opts="-l"
```

Libvirt.conf 文件的配置需要保持默认值。Libvirtd.conf 文件需要配置监听方式安全证书，修改的参数如下，SSL 安全证书的制作与配置见下一小节：

```
listen_tcp = 1          启用 TCP 协议监听
auth_unix_ro = "none"   不允许任何人连接只读的 unix 套接字
auth_unix_rw = "none"   不允许任何人连接可读写的 unix 套接字
```

4.4.4 制作 libvirtd 安全通信使用证书

dosvmd 程序需要联合 libvirtd 程序来实现虚拟机的控制，连接 libvirtd 程序需要使用 SSL 方式加密。使用 SSL 通信需要为通信的双方签发安全证书。

制作安全通信需要的证书的脚本在目录“/usr/local/bin/dosvm/cert/”中。

4.4.4.1 制作签证机构证书

如果你使用权威机构签发的安全证书，则可以直接从权威机构获得安全通信需要的各种证书。当然，你可以自己扮演权威机构，自己为自己签发证书。在给通信实体签发证书前，需要有机构证书（即 CA 证书，也叫“根证书”）。如果你已经有机构证书，则可以跳过本步骤。

制作机构证书的脚本为 gen_cacert.sh。你在数字有机体系统的终端中执行以下命令即可制作机构证书。

```
cd /usr/local/bin/dosvm/cert
./gen_cacert.sh
```

制作好的证书就放在当前目录下。公开证书的文件名称为 cacert.pem。这是一个需要分发给各个通信实体的证书。私有证书的文件名 cakey.pem，这个文件需要安全妥善的保存。

4.4.4.2 制作服务器证书

需要为每台服务器制作通信证书。你可以在数字有机体系统的终端中，在进入/usr/local/bin/dosvmd/cert 目录后，执行以下命令制作服务程序的通信证书。

```
./gen_svr_cacert.sh  服务器的通信 IP
```

该脚本的执行参数为服务器的通信的 IP 地址。如果服务器有多个网络接口，则只有其中一个接口用于内部网络通信，即服务器与服务器间的通信接口。本节设置的安全通信是服务器间的，因此需要使用的是内部网络通信接口的 IP 地址。

该程序将显示很多信息，最终提示“Signing certificate...”。如果证书制作成功，将在当前目录下生成以服务器通信 IP 为名称的目录，进入该目录，可以列出生成的证书：

```
cacert.pem  clientcert.pem  clientkey.pem  servercert.pem  serverkey.pem
```

以下是配置文件“/etc/libvirt/libvirtd.conf”中的配置情况，如果不使用缺省的证书路径，应该先删除配置文件中每行开头的“#”，然后修改为最终的证书路径。

```
#################################################################
#
# TLS x509 certificate configuration
#



# Override the default server key file path
# 服务端的私钥证书
#key_file = "/etc/pki/libvirt/private/serverkey.pem"

# Override the default server certificate file path
# 服务端的公开证书
#cert_file = "/etc/pki/libvirt/servercert.pem"

# Override the default CA certificate path
# 机构的公开证书，这个文件是前节所说的机构证书的拷贝
#ca_file = "/etc/pki/CA/cacert.pem"
```

此外，客户端的安全公开证书 clientcert.pem 需要拷贝到目录“/etc/pki/libvirt/”中；私钥证书 clientkey.pem 需要拷贝到目录“/etc/pki/libvirt/private/”中。

4.4.5 制作 https 安全通信使用的证书

为了支持安全 http 协议的访问，需要给 tomcat 服务器制作 keystore，并给用户的浏览器（如 IE）制作根证书。制作的步骤如下：

1) 在终端执行“cd /usr/local/bin/dosvm/cert/”命令，进入制作证书的目录；
2) 执行“./tomcat_cert.sh 保存目录”命令，并按照提示输入密码；
3) 执行结束后，最终获得两个文件：keystore 和 cacert.cer。

在服务端，tomcat 服务器需要在配置文件“/usr/local/tomcat/conf/ server.xml”中配置 keystore 的路径和密码：

```
<Connector SSLEnabled="true" URIEncoding="utf-8" clientAuth="false"
        keystoreFile="/usr/local/tomcat/conf/keystore" keystorePass="123456"
        maxThreads="150" port="8443" protocol="HTTP/1.1"
        scheme="https" secure="true" sslProtocol="TLS"/>
<Connector SSLEnabled="true" URIEncoding="utf-8" clientAuth="false"
        keystoreFile="/usr/local/tomcat/conf/keystore" keystorePass="123456"
        maxThreads="150"
```

```
port="8446"
protocol="org.apache.coyote.http11.Http11NioProtocol"
scheme="https" secure="true" sslProtocol="TLS"/>
```

在客户端，用户在使用 https 之前，应该先要获取到根证书“cacert.cer”，然后再把它导入到浏览器，告诉浏览器虚拟机管理网站是可信的。

4.4.6 制作 dosvmd 与管理网站安全通信使用的证书

dosvmd 是要和 WEB 通信的，为了保证通信的安全，这里需要为它们制作通信安全证书。

制作安全通信需要的证书的脚本在目录“/usr/local/bin/create_ca/”中，脚本“create_ca”同时提供制作根证书和服务证书的能力。

4.4.6.1 制作守护程序 dosvmd 的通信证书

dosvmd 程序使用的数字证书是用 OpenSSL 制作的。该证书内不仅封装了颁发证书机构的信息，还包含了证书持有者的 IP。在鉴别证书时，不仅鉴别对方的证书真伪，还鉴别对方的 IP 和证书持有者 IP 是否相同。制作方式如下：

1) 运行脚本“create_ca”，制作机构证书时，将提示输入加密证书的密码，请按照如下的提示输入并再次输入以进行确认。这个密钥是后面制作主机证书时要求输入的，因此不要忘记了。

```
开始制作cacert.pem
Generating a 1024 bit RSA private key
..++++++
.............................................++++++
writing new private key to '/usr/local/bin/create_ca/CA/private/cakey.pem'
Enter PEM pass phrase:

Country Name (2 letter code) [CN]:
State or Province Name (full name) [sichuan]:
Locality Name (eg, city) [chengdu]:
Organization Name (eg, company) [txy]:
Organizational Unit Name (eg, section) [txy]:
Common Name (eg, YOUR name) [192.168.0.100]:qyjyuanjie
Email Address [tianxinyue@126.com]:
```

2) 制作主机证书时，也将首先生成主机的私钥，并提示输入封装私钥的密码，如下所示。在输入并确认密码后，即开始输入主机的信息，这些信息条目和机构证书的相同，只是有一个可选的挑战密码和公司名称，建议直接回车略过。

```
开始制作key.pem
Generating a 1024 bit RSA private key
.........++++++
.................++++++
writing new private key to '192.168.0.100/server-key.pem'
Enter PEM pass phrase:
Verifying - Enter PEM pass phrase:
```

3) 完成主机证书信息输入后，将询问是否采用密码封装证书，建议选择是，这时将提示输入密码。该密码必须是第 6 步时输入的私钥封装密码，否则将出错。而且在配置参数时将使用该密码。

```
Enter pass phrase for 192.168.0.100/server-key.pem:
```

4) 输入主机的私钥封装密码后，将提示输入证书的私钥封装密码。该密码是在第 4 步时输入的，这里一定要输入正确。

```
开始制作cert.pem
Using configuration from ./openssl.cnf
Enter pass phrase for ./CA/private/cakey.pem:
```

5) 在正确输入后，将显示主机证书的信息，确认无误后，在下面提示中输入“y”即可完成主机证书制作。

```
Certificate is to be certified until May 20 09:53:52 2015 GMT (365 days)
Sign the certificate? [y/n]:y
```

6) 生成的证书在/usr/local/bin/create_ca/主机 IP/目录下。这里共有三个文件，他们分别是：cacert.pem 是机构证书文件，server-cert.pem 是主机的公开证书文件，server-key.pem 是主机的私钥证书文件。现在需要将它们拷贝到运行主机上。

7) 最后将上述三个文件放在配置文件制定的位置。具体的位置由配置文件“/etc/dosvmd.conf”中的“[ssl]”栏目选项决定，详细的参数介绍在后面描述。

4.4.6.2 制作管理网站 dosvm.war 的通信证书

使用 OpenSSL 工具制作的证书无法被 WEB 程序使用。WEB 程序（dosvm.war）需要 PKCS12 格式的根证书和服务证书。这里需要根据前面生成的文件制作应用于 WEB 的证书。具体的做法是把根证书的公开证书 cacert.pem 和私钥证书 cakey.pem 合并为一个格式为 PKCS12 格式的文件 cacert.p12。把服务公开证书 server-cert.pem 和私钥证书 server-key.pem 合并为 server-cert.p12。

在制作的过程中需要设置密码，请务必记住密码，在后面的配置中将会用到。制作的步骤如下：

1) 执行“./trans_ca”命令，把根证书的公开证书 cacert.pem 和私钥证书 cakey.pem 合并为一个格式为 PKCS12 格式的文件 cacert.p12，这里需要输入原来的 ssl 证书的密码，还要输入生成的 PKCS 文件的密码；

2) 执行“./trans_cert 192.168.0.100”命令，这里的 IP 地址（如“192.168.0.100”）必须和前面的 create_ca 使用的地址一致，这里的 IP 实际上可以理解为需要转换证书所在的目录。转换结束后将产生服务证书 server-cert.p12。

这里制作的 PKCS12 格式的证书 cacert.p12 和 server-cert.p12，以及证书的密码将会在 WEB 网站中使用。配置的具体方式请参考 13.2.2 章节。

4.4.7 配置守护程序 dosvmd 的参数

在每台服务器上，守护程序 dosvmd 都有一个参数配置文件，即/etc/dosvmd.conf。其中的某些参数是需要每台服务器不同的，因此必须单独配置。需要直接编辑该文件来设置参数。

可以使用 vim 来编辑它。配置文件中，用“#”符号开始的为注释行。文件的内容大致如下。

```
# Configuration file used by 'dosvmd'
#-----------------------------------------------------------------
#
# PLEASE READ THIS NOTE:
# ---------------------
# Copyright (c) 2015 - 2020, ChengDu Tian Xin Yue High-Tech CO..LDT.
# All rights reserved.
#
# You can uncomment this arguments to use default options
# MUST CONFIG:
# -----------
# (1).db
# (2).localhost
#
##############################################################################
# Here begins the real configuration file

[network]
localhost  = 192.168.2.189
min_tcp_port = 2000
max_tcp_port = 4000

[db]
db_name = dosvm
db_host  = localhost
db_user  = dba
db_password = sql
db_max_connect_count = 2

[command_thread]
max_command_thread_count = 10

[ssl]
use_ssl=0
ssl_ca=
ssl_cert=
ssl_key=
```

ssl_password=

[config]
temp_image_save_path=/tmp
share_disk_path=/raid/data
host_uuid=
libvirt_url=qemu:///system

[info]
vir_log_save_path=/var/log/

配置文件的 network 节配置网络相关的参数。各个参数的含义是：

localhost：配置宿主机的内部通信 IP 地址。宿主机间需要相互协作，因此也需要相互通信。现在只支持 IPv4 地址。

min_tcp_port：配置宿主机间相互通信时使用的 TCP 协议端口范围，该参数时范围的起始值。max_tcp_port 则配置其最大值。包括最小和最大值。注意该范围一定要有充足的端口数，建议不小于 100 个端口。

max_tcp_port：宿主机间相互通信时使用的 TCP 协议端口范围。

配置文件的 db 节配置虚拟机系统的数据库的访问参数。虚拟机系统需要使用一个共享的数据库保存公共信息。通常这个数据库建立在数字有机体数据库系统中，由数字有机体数据库系统提供透明访问、复制等功能。各个参数的含义如下：

db_name：虚拟机系统数据库的名称

db_host：数据库服务器的地址，如果本服务器也加入了数据库服务系统，则可以使用 localhost，否则需要填写数据库服务系统中的某台服务器的地址。

db_user：访问数据库的用户名。

db_password：访问数据库使用的口令。

db_max_connect_count：访问数据库的最大连接数。如果只有一个连接可能影响性能。建议在 2 到 4 个间。

配置文件的 command_thread 节只有一个参数，用于配置消息出来线程池的线程个数。根据消息通信的频繁程度设定。建议在 5 到 10 个间。

配置文件的 ssl 节用于配置安全通信的参数。本系统使用安全套接字层来加密通信数据。这里配置安全通信需要的证书等信息。证书的制作参见后面的章节。

use_ssl：是否使用安全通信，0 表示不使用，1 表示要使用。

ssl_ca：系统根证书文件的全路径名。所有服务器使用相同的根证书。

ssl_cert：本机证书文件的全路径名。每台服务器都需要一个独有的认证证书。

ssl_key：本机私钥文件的全路径名。这个文件是和证书文件对应的，必须一次生成服务器的证书文件和私钥文件。

ssl_password：访问证书文件的密码。制作证书时可以设定证书的加密密码。这里可以配置它。

配置文件的 config 节用于配置程序的运行参数。每个参数的含义如下：

temp_image_save_path：在启动虚拟机时，可能需要为虚拟机磁盘创建临时的快照。这里设定存储临时快照的路径。要注意的是，必须确保该目录所在的磁盘有足够的空间来存储这些临时快照，否则可能因无法创建临时快照而不能启动虚拟机。

share_disk_path：宿主机共享磁盘的路径。这个路径是数字有机体存储系统使用的，本系统仅仅在收集宿主机信息时用于获得宿主机的共享磁盘空间信息，如总容量和空闲容量。

host_uuid：主机的唯一标识，采用字符串形式的 UUID 格式。建议不要配置，守护程序在第一次启动时会自动生成一个全局唯一的标识，并写入该配置项。这里仅仅起到永久保存的目的。

libvirt_url=qemu:///system，访问宿主机的模拟器的 url。建议不要修改它。这是 libvirtd 默认的访问 url。

配置文件的 info 节仅有一个参数，用于配置虚拟机运行日志的保存路径。建议保存在/var/log 目录下。注意，如果需要长期保存虚拟机的运行日志，则要求目录所在的磁盘有足够的存储空间。因此，也可以建立一个单独的磁盘分区来保存虚拟运行日志。

修改配置文件后，重启起效，启动的方式如下所示。

4.5 软件的运行启动

守护程序 dosvmd 需要在每台服务器上启动，而管理网站则只需在部分服务器上启动。

可以将 dosvmd 设为自动启动的系统服务。这样在服务器启动时服务就能自动启动。设定 dosvmd 自动启动的命令为：

update-rc.d　dosvmd start

update-rc.d　dosvmd enable

如果不希望 dosmvd 自动启动，则可以使用如下的命令关闭自动启动：

update-rc.d　dosvmd stop

update-rc.d　dosvmd disable

也可以手动启动 dosvmd 服务。该程序的命令行格式为：

./dosvmd [options]

其中 options 是可选的参数，可以是“-l”和“-d”，不加参数时，默认是不记录日志并后台运行的。使用参数“-l”时，服务程序将记录运行日志到/var/log/dosvmd.log 文件中，否则服务程序不记录日志。当需要观察服务运行时，可以使用该参数。当正式运行时建议不要

使用该参数。使用参数“-d”时，服务程序将运行日志打印到屏幕上，并且不启动后台运行，当调试程序时可以使用。另外，运行“./dosvmd --usage”时可以查看该程序的使用方法；运行“./dosvmd --help”时可以获取该程序的帮助信息；运行“./dosvmd --version”时可以获取该程序的版本。

管理网站部署在 tomcat 服务中，随 tomcat 一起启动。tomcat 也可以作为自动启动服务。且设定方式和 dosvmd 的设定方式相同。不同的仅仅是将 dosvmd 替换为 tomcat7。

也可以手动启动或者停止 tomcat 服务。启动的命令为：

service tomcat7 start

停止的命令为：

service tomcat7 stop

重新启动的命令为：

service tomcat7 restart

tomcat 服务正常启动后，即可通过浏览器访问管理网站。系统管理员的账号为“sysadm”，初始口令为“123456”，建议一旦正式使用，应该立即修改口令，以保证系统安全。

5. 用户管理

5.1 用户分类介绍

在数字有机体管理网站上，用户被分为系统管理员和项目管理员两类。一个系统中只有一个系统管理员，它负责完成系统级的管理工作。同时，也由系统管理员管理项目管理员账号。项目管理员管理具体的某个项目。一个项目管理员账号只能管理一个项目，但是一个项目可以有多个项目管理员。这里的项目是一个虚的概念，可以看做是一组虚拟机应用的集合。

系统管理员管理的模块包括用户管理、项目管理、存储池管理、网络管理、安全对象管理、宿主机群管理、操作日志管理和系统参数配置。

项目管理员管理的模块包括虚拟机管理、虚拟机机群、存储卷管理、镜像管理、网络管理、资源使用情况、运行监控和查询、个人信息管理。

用户管理页面结构如图 5-1 所示：

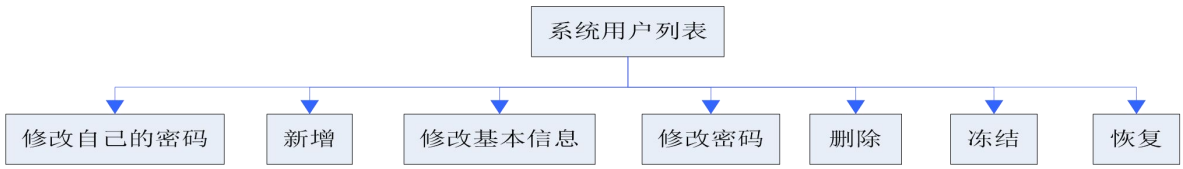


图 5-1：用户管理页面的结构

5.2 登录管理系统

对系统的操作大都要在管理网站中完成。在启动数字有机体管理网站的服务后，就可以登录管理系统。

默认地，数字有机体系统中的 tomcat 在 8080 端口上提供服务。数字有机体虚拟机管理系统的管理网站部署在 tomcat 中，应用名称为 dosvm。因此，在浏览器里，可以采用如下地址访问数字有机体管理网站：

http://管理网站的服务器 IP 地址:8080/dosvm/

如果已配置 https 访问的安全证书，并且已将浏览器设定了该管理网站所用证书为可信证书，还可以采用以下的安全连接来访问：

https://管理网站的服务器 IP 地址:8080/dosvm/

如果系统安装配置正确，将在浏览器里看到如图 5-2 所示的登录界面。图中管理网站的服务器 IP 地址为 192.168.2.189。服务端口仍然是 8080，login.jsp 是自动跳转的页面。

注意：目前支持的浏览器是 IE 和火狐，其它浏览器未经测试。

系统安装后，初始的系统管理员账号为“sysadm”，口令为“123456”。系统管理员的账号名称不能修改。因此，系统管理员在用初始密码登录系统后，建议立即修改密码，以保证系统管理员账号的安全。

在登录界面上输入登录账号和密码后，点击登录即可进入管理系统。

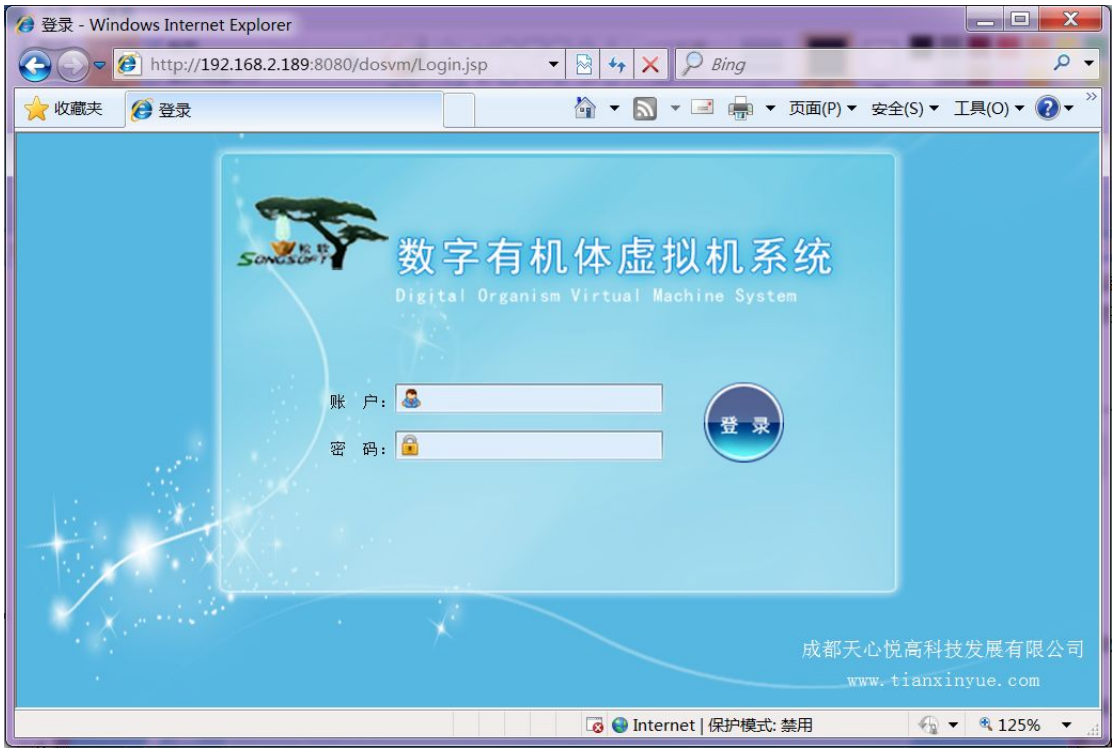


图 5- 2：登录数字有机体虚拟机系统管理网站的界面

5.3 系统管理流程概述

在应用系统时，首先应由系统管理员登录系统进行系统配置。系统配置的工作流程大致如下：

1）建立项目管理员，以便让他管理具体的项目；

2）管理系统的服务器，建立必要的宿主机群；

3）为虚拟机应用建立项目，为项目配置管理员，设置配额；

4）根据网络部署方案配置子网和安全组，并把子网分配给项目；

5）为访问外部资源需要的密码、证书等配置安全对象；

6）配置系统可用的存储池，并把存储池分配给项目使用。

在完成上述配置后，项目的管理员即可登录管理系统，管理自己的项目。项目管理员是运行和控制虚拟机的具体人员，系统管理员不直接管理和运行虚拟机。

要注意的是：必须为项目分配资源，否则项目将没有资源可以使用。主要的资源是存储池、网络子网。

项目管理配置虚拟机的流程大致如下：

1）如果虚拟机要使用镜像文件（用于模拟自读磁盘或者光驱），则先上传或者建立镜像文件。

2）如果虚拟机要使用存储卷，则在项目可用的存储池中创建需要的卷；

3）如果虚拟机要使用网络，则根据网络部署规划，利用可用的子网构建虚拟网络；

4）为虚拟机实例增加虚拟机。

现在可以直接控制和运行虚拟机了。如果需要监控虚拟机的运行情况，可用在“运行情况与查询”模块中进行。如果要了解项目资源的使用情况，可用在“资源使用情况”模块查看。

5.4 系统管理员管理用户

用系统管理员账户登录系统后，可以看到如图 5-2 所示的界面。

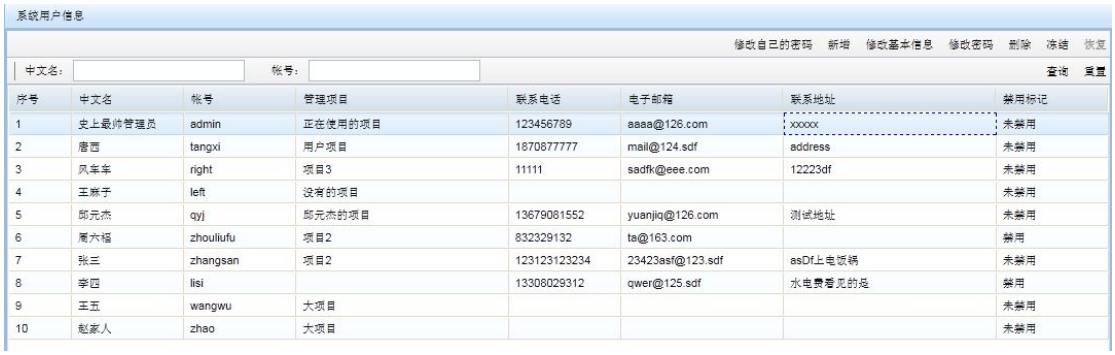
图 5- 3：系统管理员的管理网站主界面

该界面的上部是导航栏，系统管理员的主要工作版块都在这里，可以点击相应按钮进入不同的管理模块。

该界面的左侧为二级导航栏，用于在同一个版块下切换不同的子模块。用户管理版块下可以管理系统的项目管理员账户，另外有关于本软件的说明。

5.4.1 系统用户信息

本系统的用户即是项目管理员。用户管理版块的首页面如图 5-4 所示。



序号	中文名	帐号	管理项目	联系电话	电子邮箱	联系地址	禁用标记
1	史上最帅管理员	admin	正在使用的项目	123456789	aaaa@126.com	xxxxx	未禁用
2	唐西	tangxi	用户项目	1870877777	mail@124.sdf	address	未禁用
3	风车车	right	项目3	11111	sadfk@eee.com	12223df	未禁用
4	王麻子	left	没有的项目				未禁用
5	邱元杰	qyj	邱元杰的项目	13679081552	yuanjiq@126.com	测试地址	未禁用
6	周六福	zhouliufu	项目2	832329132	ta@163.com		禁用
7	张三	zhangsan	项目2	123123123234	23423asf@123.sdf	asDf上电饭锅	未禁用
8	李四	lisi		13308029312	qwer@125.sdf	水电费看见的是	禁用
9	王五	wangwu	大项目				未禁用
10	赵家人	zhao	大项目				未禁用

图 5- 4：用户管理版块的界面

系统用户信息主页显示了系统用户（项目管理员）的信息，包含中文名、账号、管理项目名称、联系电话、禁用标记等信息。

在该页面上可以根据项目管理员的中文名称或者账户查询。系统支持模糊匹配，以便快速检索需要的项目管理员。可以只输入中文名称或者账户，点击“查询”按钮即可进行查询。如果要清除查询设置，需要先点击“重置”按钮，再点击一次“查询”按钮。重置操作仅仅是清空了已经输入的查询条件，但并没有进行查询。

5.4.2 修改系统管理员密码

在页面列表上部的工具栏上点击“修改自己的密码”按钮，弹出修改系统管理员密码对话框如图 5-5 所示。

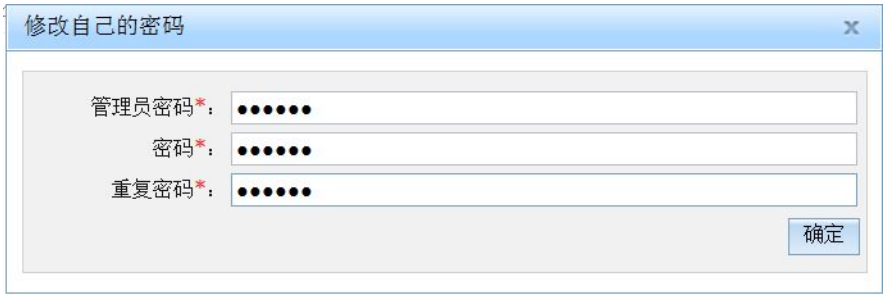


图 5- 5：修改系统管理员的密码

首先输入系统管理员原来的密码，然后输入新密码，再确认输入新密码。密码的长度不超过 16 位。

5.4.3 新增项目管理员

在页面列表上部的工具栏上点击“新增”按钮，弹出新增项目管理员对话框，如图 5-6 所示。

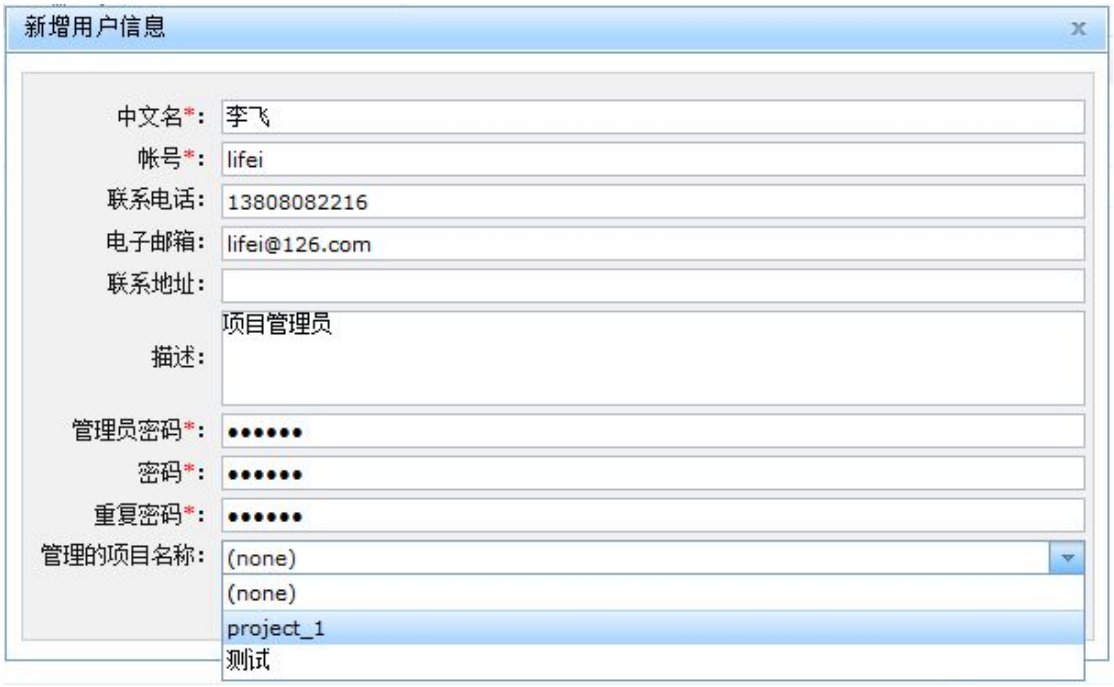


图 5- 6：新增用户的界面

新增用户信息页面可输入中文名、账号、联系电话、电子邮箱、联系地址、描述、系统管理员密码和新增的用户密码。为了避免管理员登陆的终端被他人盗用，这里要求输入管理

员的密码，以便进行验证。所有加星号的条目都是必须输入的，新增用户时可以设定该用户所属的项目，也可暂不设置，之后再在项目管理中设置。

5.4.4 修改基本信息

在用户信息列表中选择一个项目管理员，然后在列表上部的工具栏上点击“修改基本信息”按钮，弹出修改该项目管理员基本信息的对话框，如图 5-7 所示。

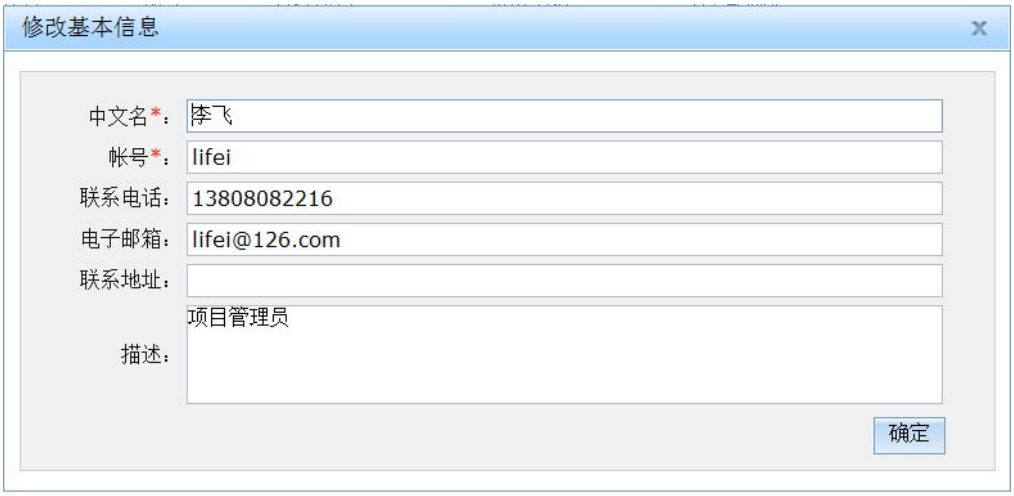


图 5- 7：修改项目管理员基本信息的界面

在该界面上可以修改用户的中文名、账号、联系电话、电子邮箱、联系地址和描述信息。项目管理员的密码不再该界面上设定。

5.4.5 修改项目管理员密码

先在用户信息列表中选择一个项目管理员，然后在列表上部的工具栏上点击“修改密码”按钮，弹出修改项目管理员密码对话框，如图 5-8 所示。

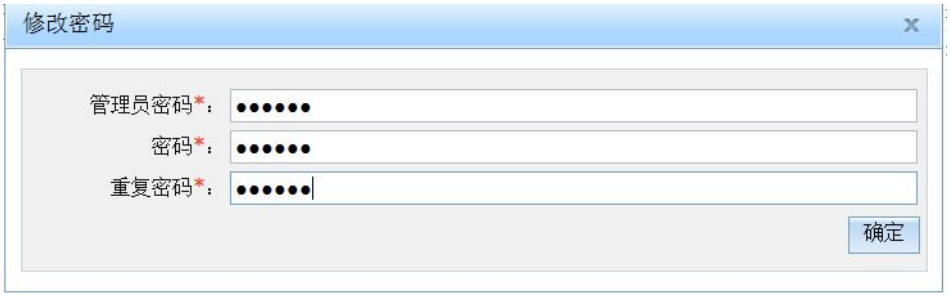


图 5- 8：修改项目管理员密码的界面

为了防止他人盗用系统管理员的登录终端，这里需要输入系统管理员的密码（在管理员密码输入框中），系统将检查系统管理员的密码是否正确。后面两个密码输入框用于输入项目管理员的新密码。密码的长度不超过 16 位。

5.4.6 删除用户

先在用户信息列表中选择一个项目管理员，然后在列表上部的工具栏上点击“删除”按钮，弹出输入系统管理员密码对话框，如图 5-9 所示，输入密码后点击确定将删除该项目管理员。如果输入的系统管理员的密码是错误的，系统将拒绝操作。这对上述各个需要输入系统管理员密码的操作都是一样的。

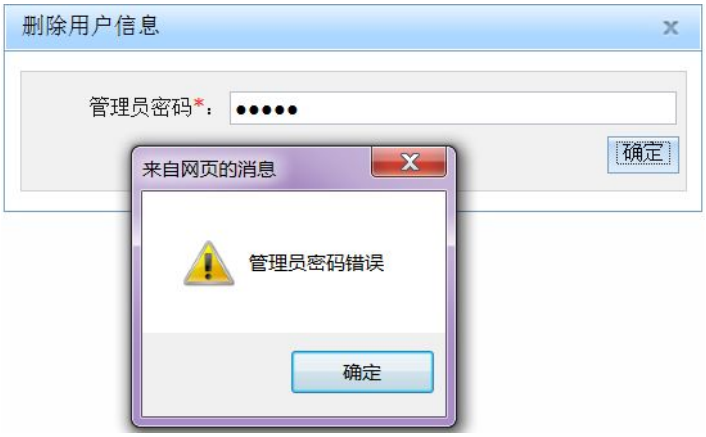


图 5- 9：删除项目管理员的确认界面

5.4.7 冻结/恢复用户

先在用户信息列表中选择一个项目管理员，然后在列表上部的工具栏上点击“冻结”或“恢复”按钮，将冻结或恢复该项目管理员的账号。冻结项目管理员后，该项目管理员将不能再登录系统。不过，如果该项目管理员已经登录，并不立即将其踢出系统，因此在该项目管理员未退出前，他仍然可以进行操作。

如果要恢复某个项目管理员的登录，则可以进行恢复操作。

5.5项目管理员个人信息管理

项目管理员登录系统后不能管理其他账号，只能对自己的账号进行管理。

5.5.1 个人信息管理主页面

项目管理员个人信息管理页面结构如图 5-10 所示：

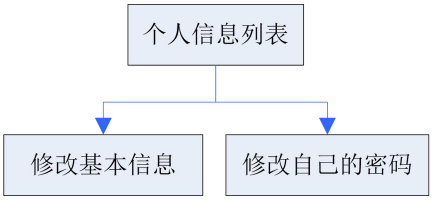


图 5- 10：项目管理员个人信息管理页面结构

项目管理员在登录系统后，可以点击顶端导航栏中的“个人信息管理”按钮，进入如图 5-11 所示的个人信息管理页面。点击个人信息列表中的行，将激活“修改基本信息”和“修改自己的密码”按钮。

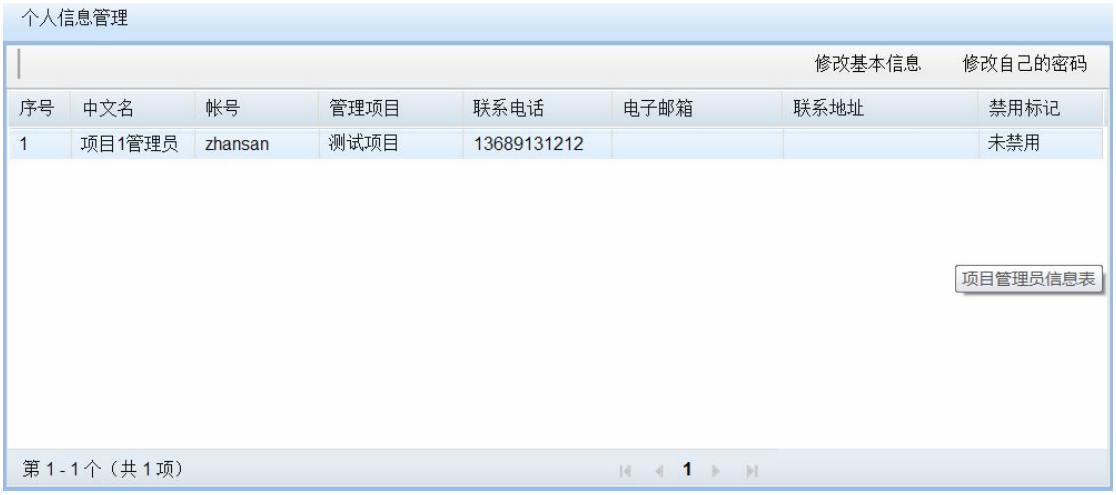


图 5- 11：项目管理员个人信息管理界面

个人信息管理页面显示了当前项目管理员的信息，如中文名、账号、管理的项目名称、联系电话、禁用标记等信息。

5.5.2 修改基本信息

在个人信息列表上面的工具栏上点击“修改基本信息”按钮，弹出修改基本信息对话框，如图 5-11 所示。

这里可修改当前管理员的中文名、联系电话、电子邮箱、联系地址和描述信息。账号是不能修改的。

5.5.3 修改密码

在上部工具栏上点击“修改基本信息”按钮，弹出修改自己的密码对话框，如图 5-12 所示。

图 5- 12：项目管理员修改个人密码的界面

修改密码先输入新密码，然后确认输入新密码。密码的长度不超过 16 位。

6. 服务器管理

6.1 概念介绍

宿主机是虚拟机运行的物理机，即系统中的服务器。虚拟机运行在宿主机中，使用宿主机的资源。虚拟机可以运行在不同的宿主机中，只要宿主机能够为虚拟机提供充足的资源。

宿主机群是一台及以上的宿主机组成的集合。虚拟机在线迁移要求源宿主机和目的宿主机具有相同或者相似的配置。尤其是要求宿主机的 CPU 类型相同，外部网络的环境相同。否则，虚拟机在线迁移可能失败，甚至迁移成功后虚拟机无法正常运行。因此，常常将配置相同或者相似的服务器看做一个集合，组织为一个宿主机群。这是宿主机群产生的一个原因。另外，当系统分散分布式部署时，各个部署点的网络环境不同。而业务系统有可能希望运行在固定的某些或者某个部署点，这时也需要将宿主机组织为宿主机群。

虚拟机可以在属于同一宿主机群的、基本配置相同的宿主机间迁移。或者说，项目管理员可以配置虚拟机可以运行的宿主机群。

6.2 管理员管理服务器

6.2.1 宿主机群管理

系统管理员登陆后，点击上部导航栏的“宿主机群管理”，进入宿主机群管理界面，如图 6-1 所示。

图 6- 1：宿主机群管理界面

界面上部是导航栏，左边是二级导航栏，包括宿主机群信息和宿主机信息两个子模块，分别点击即可进入相应的子模块。

宿主机群管理页面结构如图 6-2 所示。

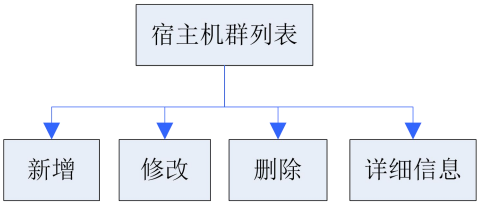


图 6- 2：宿主机群管理的页面结构

6.2.1.1 宿主机群信息主页面

点击宿主机群管理界面左边的二级导航栏的“宿主机群信息”，打开宿主机群信息页面，如图 6-3 所示。

图 6- 3：宿主机群信息页面

宿主机群信息页面上部为操作按钮和查询输入栏；中部为宿主机群信息列表；底部为宿主机群信息列表分页显示栏。

宿主机群信息主页面列表显示了宿主机群名称和描述信息。

在宿主机群信息列表上部的工具栏内的查询输入框中，输入宿主机群的部分模糊匹配名称进行查询。

6.2.1.2 新增宿主机群

单击宿主机群主页列表上部工具栏上的“新增”按钮，弹出新增宿主机群对话框，如图 6-4 所示。

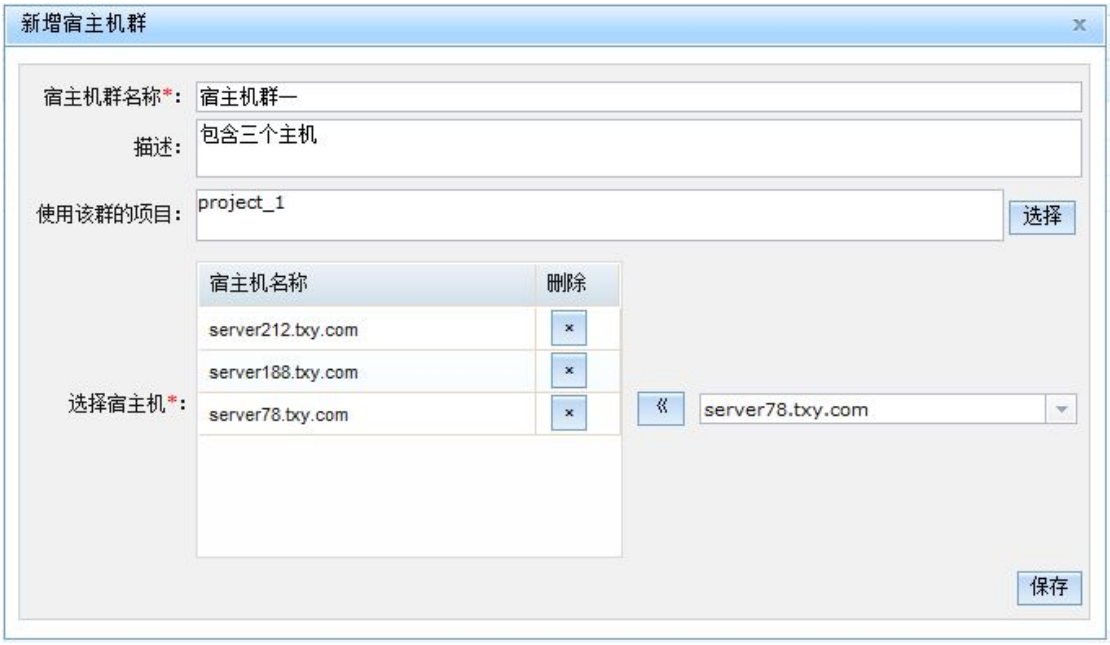


图 6- 4：新增宿主机群

新增宿主机群可输入宿主机群名称和描述。在右下角的宿主机下拉框中选择一个宿主机，然后点击左旁的“ 《 ”按钮，可将选择的宿主机添加到左边的宿主机列表中。单击宿主机列表的“X”按钮，可将宿主机从列表（宿主机群）中删除。

新增宿主机群信息输入完毕后，点击右下角的“保存”按钮，即可创建一个宿主机群。

6.2.1.3 修改宿主机群

在宿主机群主页列表中选择一个宿主机群，在上部工具栏上单击“修改”按钮，弹出修改宿主机群对话框，如图 6-5 所示。

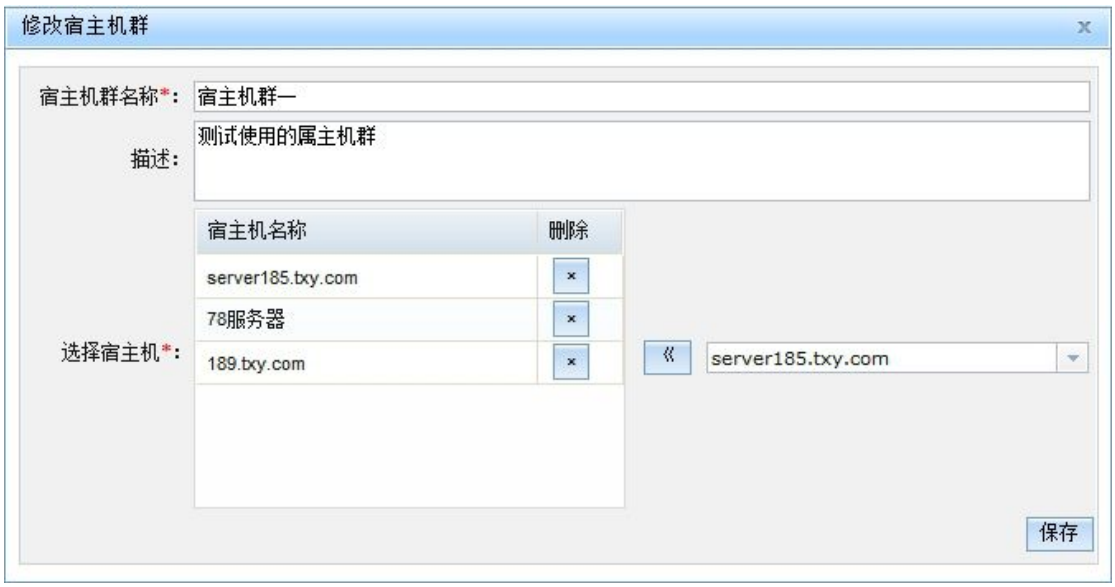
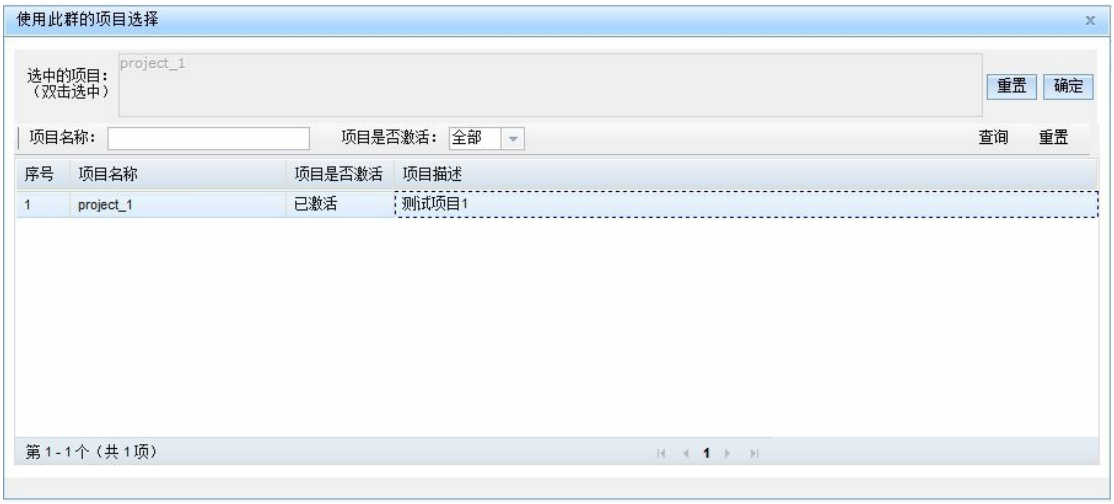


图 6- 5：修改宿主机群信息

新增宿主机群可输入宿主机群名称和描述。在右下角的宿主机下拉框中选择一个宿主机，然后点击左旁的“ 《 ”按钮，可将选择的宿主机添加到左边的宿主机列表中。单击宿

主机列表的“X”按钮，可将宿主机从列表（宿主机群）中删除。

单击“选择按钮”，弹出使用该宿主机群的项目选择对话框，如图 6-5 所示。在下部的项目列表里选择并双击，即将该项目如 project_1 添加到上部的选中的项目列表中，然后点击“确定”按钮即可。



新增宿主机群信息输入完毕后，点击右下角的“保存”按钮，即可创建一个宿主机群。

6.2.1.4 删除宿主机群

在宿主机群主页列表中选择一个宿主机群，在上部工具栏上单击“删除”按钮，弹出确认删除按钮如图 6-6 所示，点击“确定”即可删除宿主机群。如果宿主机群已经分配给了项目，或者宿主机群正作为某个虚拟机实例的运行范围，则删除它将报失败。只要没有使用的宿主机群才能被删除。

图 6- 6：宿主机群删除确认

6.2.1.5 宿主机群详细信息

在宿主机群主页面列表中选择一个宿主机群，在上部工具栏上单击“详细信息”按钮，弹出宿主机群详细信息对话框，如图 6-7 所示。

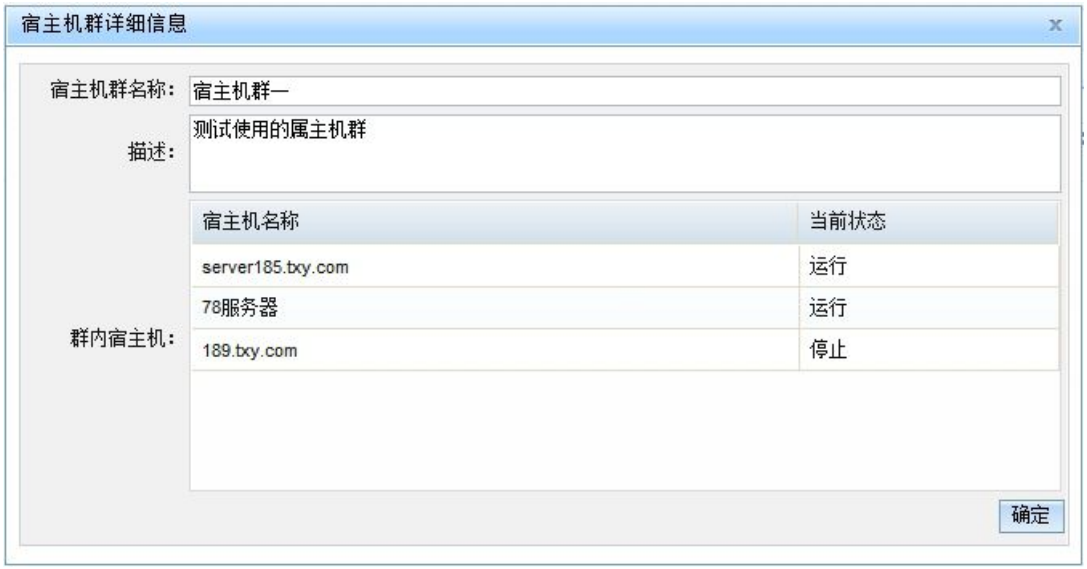


图 6- 7：宿主机群详细信息

详细信息页面显示了宿主机群名称、描述和群内宿主机列表。列表显示了宿主机名称和宿主机当前状态信息。

6.2.2 宿主机管理

管理系统的宿主机信息由守护程序 dosvmd 主动注册。当 dosvmd 启动时，它会检查系统中是否已经有自己的注册信息，如果没有则在系统中增加宿主机记录。要注意的是，dosvmd 以宿主机的 UUID 作为标记。如果修改了配置文件中主机的 UUID，则程序将自动新增加一个。这可能导致错误。因此，切勿随意修改宿主机的 UUID 配置。如果出现了同一台宿主机有多条记录，则需要手动删除它们。

宿主机管理页面结构如图 6-8 所示。

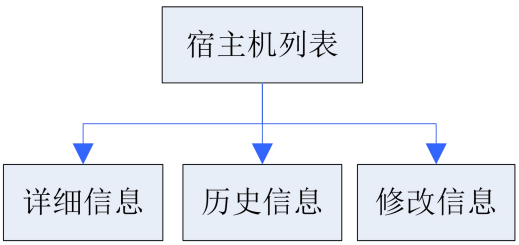


图 6- 8：宿主机管理页面结构

6.2.2.1 宿主机信息

点击宿主机群管理界面左边的二级导航栏的“宿主机信息”，打开宿主机信息页面如图 6-9 所示。

图 6- 9：宿主机信息

宿主机信息页面上部为工具栏和查询输入栏；中部为宿主机信息列表；底部为宿主机信息列表分页显示栏。

宿主机信息主页面列表显示了宿主机静态和动态信息，包括 CPU 核数、内存总数、最大 VCPU 数、当前状态、当前 CPU 负载、当前内存负载、当前磁盘负载、当前网络负载。

“当前状态”的值为运行或停止；

“当前 CPU 负载”表示 CPU 使用时间占 CPU 总时间的比例；

“当前内存负载”表示内存已使用数量占总内存数的比例；

“当前磁盘负载”表示磁盘已使用容量占磁盘总容量的比例；

“当前网络负载”的单位为 MB/S，是宿主机上所有网卡发送和接收的字节速率的总和。

在宿主机信息列表上部的工具栏的查询输入栏中，可输入宿主机模糊匹配条件进行查询。

6.2.2.2 详细信息

在宿主机信息列表中选择一个宿主机，点击上部工具栏的“详细信息”按钮，弹出宿主机详细信息页面，如图 6-10 所示。

图 6- 10：宿主机详细信息

宿主机详细信息页面最上部是宿主机动态信息刷新时间设置输入栏；页面中部显示了宿

主机静态信息，如宿主机名、CPU 型号、CPU 核数、允许最大 VCPU 数等信息；下部是宿主机动态信息折线图，包括 CPU、MEMERY、DISK、NET 的动态信息。

在上部宿主机动态信息刷新时间输入栏中设置时间后，点击右边设置按钮，若设置成功，则弹出设置成功确认对话框，如图 6-11 所示。

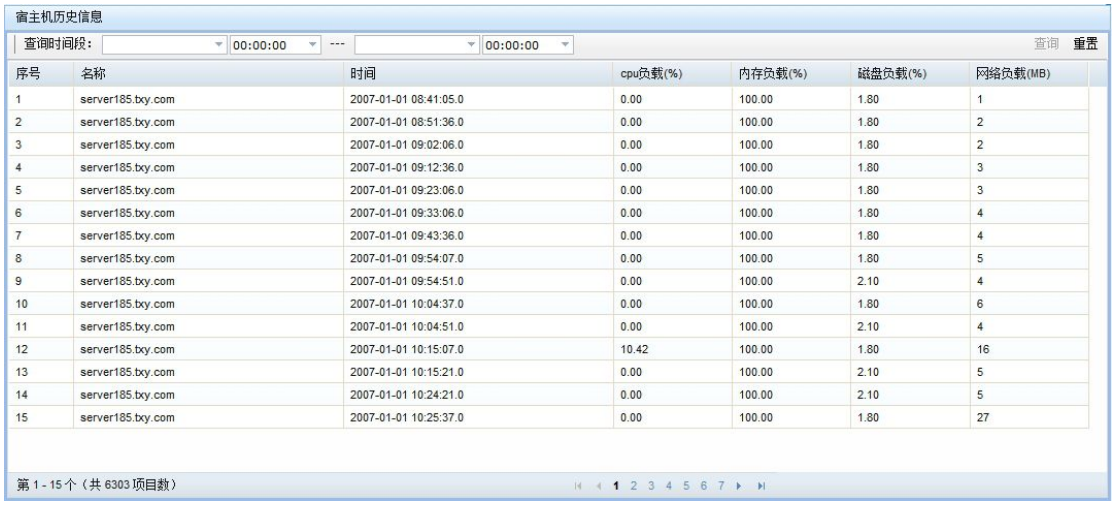
图 6- 11：宿主机动态信息刷新时间设置成功

宿主机动态信息需要周期性的从系统中收集，加上绘图需要的时间开销，因此页面弹出后需要等待数据加载。其后，页面按照设定的间隔获取新数据，并刷新页面。建议刷新的时间间隔不要太少（不小于 5 秒），否则可能占用过多的系统资源。

6.2.2.3 历史信息

系统以较大的周期保存宿主机运行的历史记录。在需要时，可以查询某段时间内宿主机的运行记录。当系统长时间运行时，系统中可能累计大量的历史记录。在数据太多时，可以清除过于陈旧的记录。如果没有查询到宿主机在某段时间的运行记录，则可能是宿主机在那个时段没有运行，或者历史记录被清除了。

在宿主机信息列表中选择一个宿主机，点击上部工具栏的“历史信息”按钮，进入宿主机历史信息页面，如图 6-12 所示。


宿主机历史信息

查询时间段： 00:00:00 --- 00:00:00 查询 重置

序号	名称	时间	cpu负载(%)	内存负载(%)	磁盘负载(%)	网络负载(MB)
1	server185.txy.com	2007-01-01 08:41:05.0	0.00	100.00	1.80	1
2	server185.txy.com	2007-01-01 08:51:36.0	0.00	100.00	1.80	2
3	server185.txy.com	2007-01-01 09:02:06.0	0.00	100.00	1.80	2
4	server185.txy.com	2007-01-01 09:12:36.0	0.00	100.00	1.80	3
5	server185.txy.com	2007-01-01 09:23:06.0	0.00	100.00	1.80	3
6	server185.txy.com	2007-01-01 09:33:06.0	0.00	100.00	1.80	4
7	server185.txy.com	2007-01-01 09:43:36.0	0.00	100.00	1.80	4
8	server185.txy.com	2007-01-01 09:54:07.0	0.00	100.00	1.80	5
9	server185.txy.com	2007-01-01 09:54:51.0	0.00	100.00	2.10	4
10	server185.txy.com	2007-01-01 10:04:37.0	0.00	100.00	1.80	6
11	server185.txy.com	2007-01-01 10:04:51.0	0.00	100.00	2.10	4
12	server185.txy.com	2007-01-01 10:15:07.0	10.42	100.00	1.80	16
13	server185.txy.com	2007-01-01 10:15:21.0	0.00	100.00	2.10	5
14	server185.txy.com	2007-01-01 10:24:21.0	0.00	100.00	2.10	5
15	server185.txy.com	2007-01-01 10:25:37.0	0.00	100.00	1.80	27

第 1 - 15 个（共 6303 项目数） 1 2 3 4 5 6 7

图 6- 12：宿主机历史记录查询

宿主机历史动态信息页面列表中显示了宿主机历史动态信息，包括宿主机名称、信息收集时间、CPU 负载、内存负载、磁盘负载和网络负载。

在宿主机历史信息列表上部的工具栏内的查询输入框中，输入时间段进行查询。

6.2.2.4 修改宿主机信息

在宿主机信息列表中选择一个宿主机，点击上部工具栏的“修改信息”按钮，弹出宿主机修改信息对话框，如图 6-13 所示。

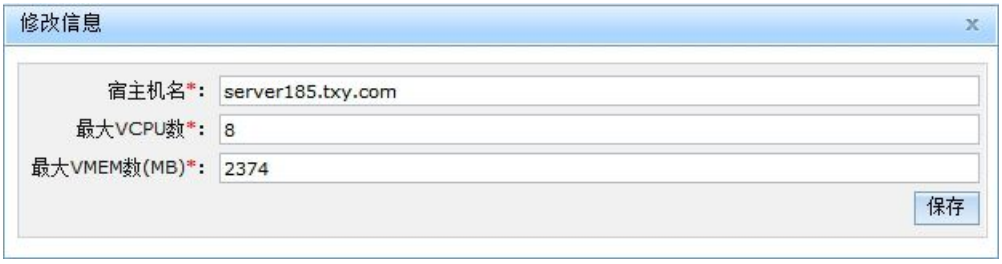
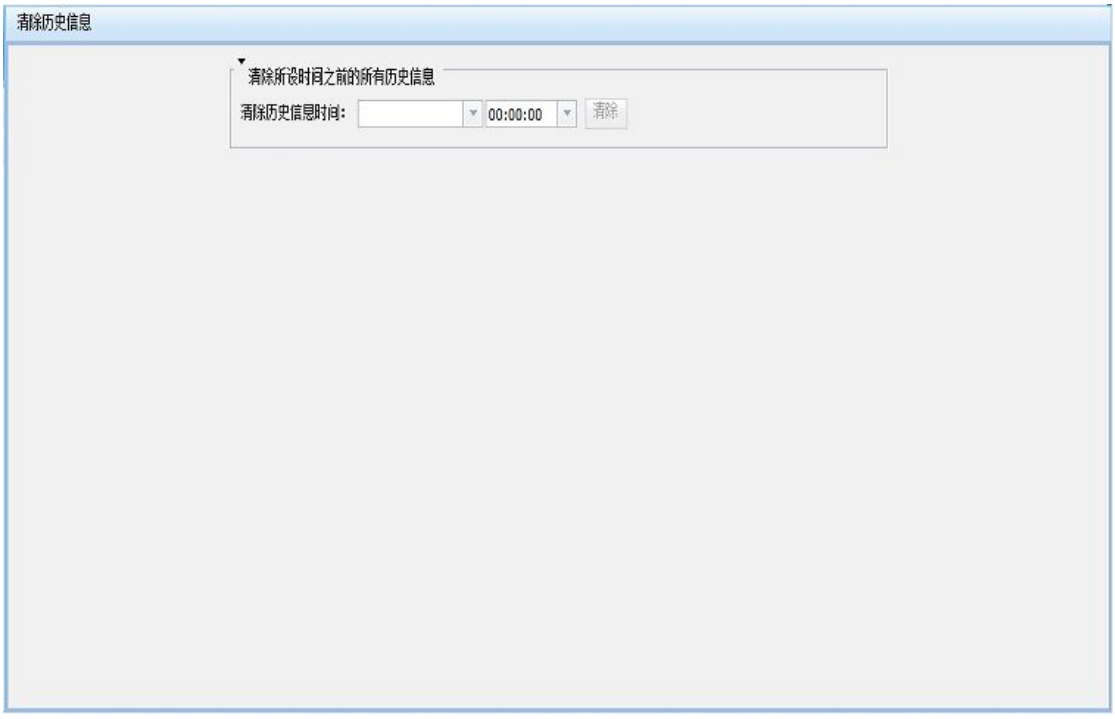


图 6- 13：修改宿主机信息

宿主机信息修改可修改宿主机名、最大 VCPU 数和最大 VMEM 数。最大 VMEM 数的最大值为宿主机物理内存数减去 512MB，此 512MB 内存为宿主机系统运行预留。

6.2.2.5 清除历史信息

点击上部工具栏的“清除历史信息”按钮，进入清除历史信息页面，如图 6-13 所示。



在历史信息页面日期控件设定日期，在时间控件设定时间后，点击清除按钮，即可清除该设定的时间点之前的宿主机历史信息。

7. 项目管理

7.1什么是项目

项目是组织和管理虚拟机运行的基础，是业务开展的基础。可以将某项业务看作一个项目。开展业务需要的资源，包括虚拟机，都在一个项目内分配和组织。为了统一管理，任何虚拟机都必须属于某个项目。即使系统仅为单个业务服务，也需要建立一个项目。

项目由系统管理员建立，并为项目分配各种资源，如虚拟机运行配额、存储池、子网、宿主机群等。系统管理员再为项目建立和分配项目管理员。由项目管理员管理项目内部事务。一个项目可配置多个项目管理员，一个项目管理员只能管理一个项目。

项目管理员利用项目资源开展业务，在虚拟机管理中新增、修改虚拟机实例，及使用该虚拟机实例创建和运行虚拟机。

7.2系统管理员管理项目

7.2.1 项目信息配置

项目信息配置页面结构如图 7-1 所示。

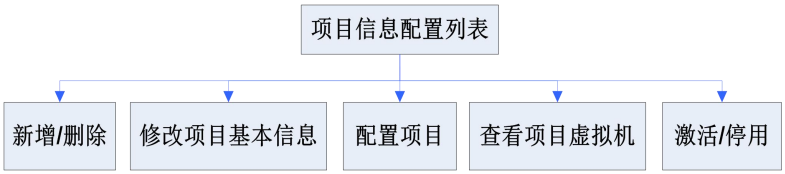


图 7- 1：项目信息配置页面结构

系统管理员登陆系统后，点击上部导航栏的“项目管理”，然后点击左边二级导航栏的“项目信息配置”，进入项目信息配置的页面如图 7-2 所示。

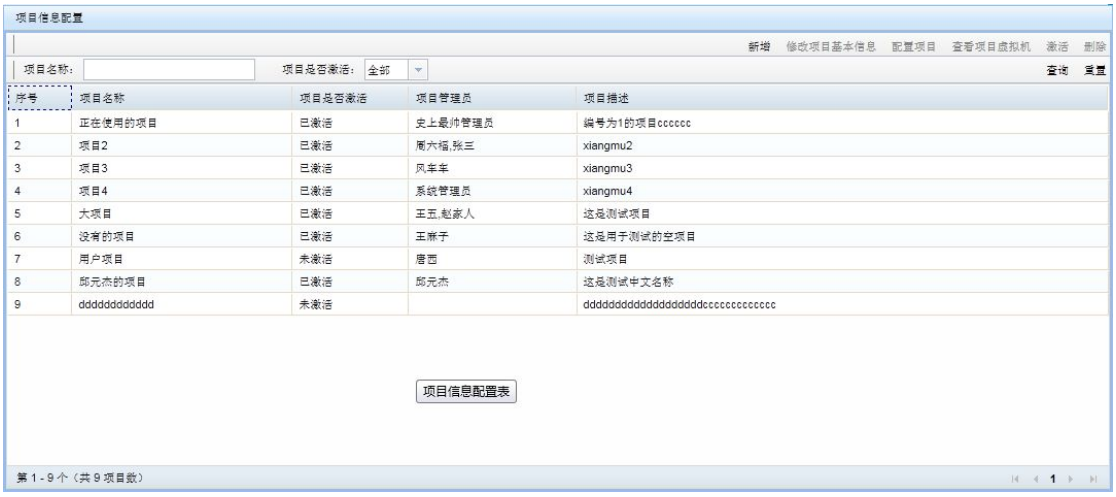


图 7- 2：项目信息配置信息

项目信息配置页面上部为操作按钮和查询输入栏；中部为项目信息配置列表；底部为项目信息配置列表分页显示栏。

在项目信息配置列表上部的工具栏的查询输入框中，输入项目的模糊匹配条件进行查询。

7.2.1.1 新增和删除项目

单击上部工具栏“新增”按钮，弹出新增项目对话框，如图 7-3 所示。

图 7- 3：新增项目

新增项目时，需在项目组下拉框中选择项目组，输入项目名称和描述，点击确定即可创建项目。如果还没有建立项目，请先到“项目组管理”模块建立需要的项目组。

选中要删除的项目，点击“删除”按钮，确定后即可删除项目。删除项目需谨慎，删除项目将删除与该被删除项目的相关信息。

7.2.1.2 修改项目基本信息

在项目信息配置主页面列表中选择一个已建立的项目，单击工具栏上的“修改项目基本信息”按钮，弹出修改项目基本信息对话框，如图 7-4 所示。

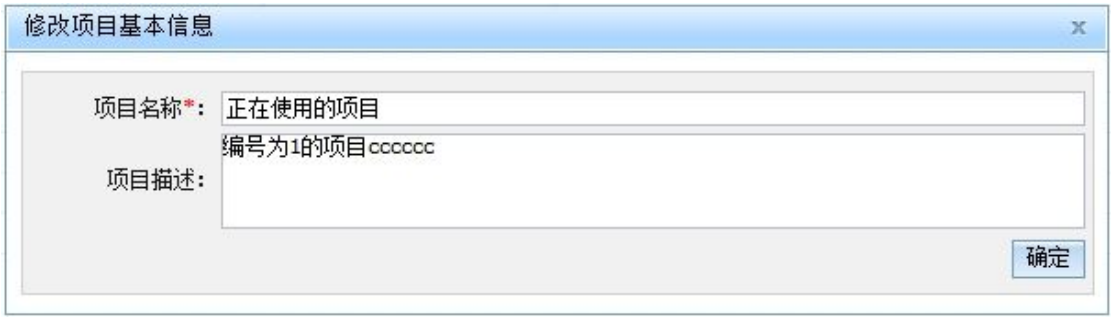


图 7- 4：修改项目基本信息

修改项目信息时，可修改项目的名称和描述。项目所属的项目组在这里无法修改。不过可以在“项目组管理”模块中修改项目所属的项目组，或者将项目加到多个项目组。

7.2.1.3 项目配置

项目不仅需要激活，还需要为其配置资源。一个没有资源的项目难以运行。在项目信息配置主页面列表中选择一个已建立的项目，点击工具栏上的“配置”按钮，弹出项目配置对话框，如图 7-5 所示。

项目需要配置的内容包括以下方面：

1）项目资源使用的配额信息，尤其是 CPU、内存和存储等资源的可用量。

2）项目的管理员：为项目指定管理员。一个项目可以有多个管理员。

3）存储池：项目可以使用的存储池。一个项目可以使用多个存储池。

4）宿主机群：项目可以使用的宿主机群。一个项目可以使用多个宿主机群。

图 7- 5：配置项目的配额信息

项目的配额信息包括项目的最大虚拟机实例数、最大 VCPU 数、最大内存数、最大磁盘容量、允许项目管理员上传的最大镜像文件总容量、允许项目管理员上传的最大镜像文件数、最大磁盘快照数。这些参数将限制项目可以使用的资源量。

配置项目的项目管理员信息页面如图 7-6 所示。

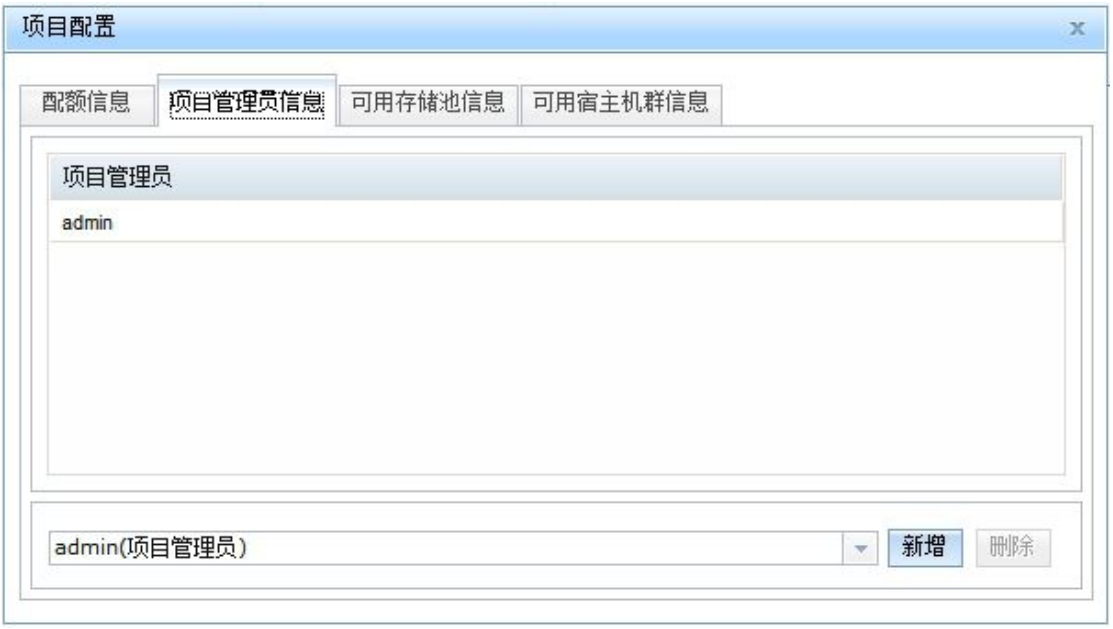


图 7- 6：配置项目的项目管理员信息

在页面底部的下拉框中选择一个项目管理员，点击“新增”按钮，将该项目管理员添加到项目管理员信息列表中。其中，一个项目管理员只能管理一个项目，不能同时管理多个项目，但是一个项目可以拥有多个管理员。

在项目管理员信息列表中选择一个项目管理员，点击右下角删除按钮，可将该项目管理员从项目管理员信息列表中删除。

配置项目的可用存储池信息页面如图 7-7 所示。

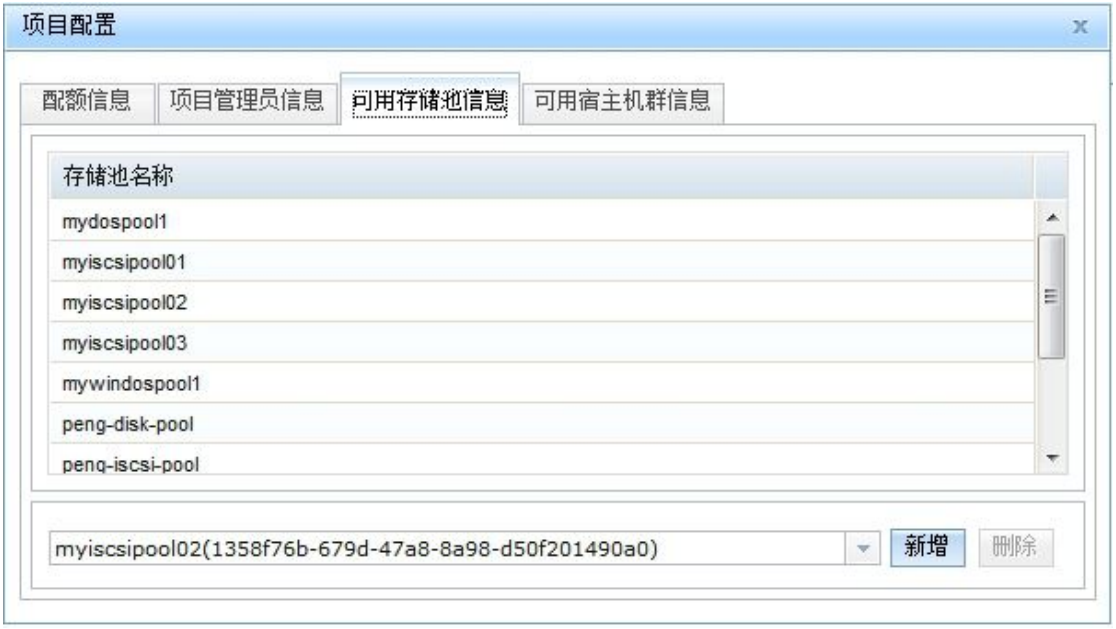


图 7- 7：配置项目的可用存储池信息

在页面底部的下拉框中选择一个存储池，点击“新增”按钮，将该存储池添加到项目可用存储池列表中；在项目可用存储池列表中选择一个存储池，点击右下角“删除”按钮，可将该存储池从项目可用存储池列表中删除。

注意：一个存储池可被多个项目共享，一个项目可有多个存储池。

配置项目的可用宿主机群信息页面如图 7-8 所示。

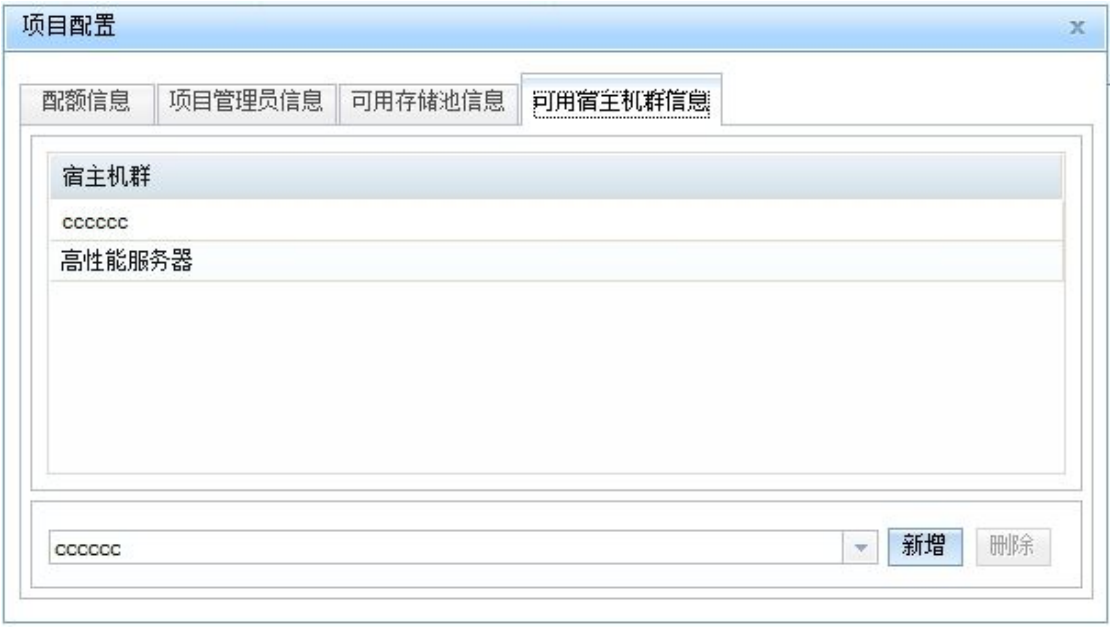


图 7- 8：配置项目的可用宿主机群信息

在页面底部的下拉框中选择一个宿主机群，然后点击“新增”按钮，即将该宿主机群添加到项目可用宿主机群列表中；在项目可用宿主机群列表中选择一个宿主机群，点击右下角“删除”按钮，可将该宿主机群从项目可用宿主机群列表中删除。

一个宿主机群可用于多个项目，一个项目可有多个宿主机群。

7.2.1.4 查看虚拟机

项目信息配置主页列表中选择一个已建立的项目，点击工具栏上的“查看虚拟机”按钮，弹出查看项目虚拟机信息对话框，如图 7-9 所示。如果不选择项目，直接点击“查看虚拟机”按钮，则是查看系统内的所有虚拟机。

图 7- 9：查看项目虚拟机信息

查看项目虚拟机对话框上部为查询输入栏；中部为项目虚拟机列表；底部为项目虚拟机

列表分页显示栏。

项目虚拟机信息列表中显示了该项目建立的虚拟机实例和使用该虚拟机实例创建的虚拟机信息。

表格中的列“是否禁用”是指虚拟机所属的实例是否禁用，已禁用的虚拟机不能再使用；

“运行主机”指虚拟机运行的宿主机地址，VNC 客户端程序远程连接时要使用，处于运行中的虚拟机总会显示此值。如果此值显示为空，表示未处于运行状态；

“访问端口”指 VNC 客户端程序要远程访问该虚拟机时，连接使用的端口号。

在项目虚拟机列表上部的工具栏内的查询输入框中，输入项目虚拟机的模糊匹配条件进行查询。

7.2.1.5 项目激活/停用

项目信息配置主页列表中选择一个已建立的项目，点击工具栏上“激活”/“停用”按钮，可激活或停用项目。项目停用后，该项目的管理员就不能对该项目进行管理，直到项目被重新激活。

7.2.2 项目组管理

项目组管理页面结构如图 7-10 所示。

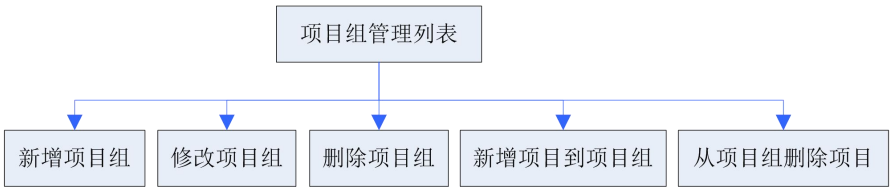


图 7- 10：项目组管理页面结构

7.2.2.1 项目组管理主页面

系统管理员登录系统后，点击上部导航栏的“项目管理”，然后点击左边二级导航栏的“项目组管理”，进入项目组管理的页面，如图 7-11 所示。

图 7- 11：项目组管理信息

项目组管理主页面上部为操作按钮和查询输入栏，中部为项目组信息列表，底部为项目组信息列表分页显示栏。

项目组管理主页面列表显示了项目组名称、该项目组包含的项目名称、项目文件保存的目录和描述信息。如果一个项目组有多个项目，则该项目组有多条记录。

在项目组信息列表上部的工具栏内的查询输入框中，输入项目组的模糊匹配条件进行查询。

7.2.2.2 新增项目组

在项目组管理主页面列表上部的工具栏上单击“新增项目组”按钮，弹出新增项目组对话框，如图 7-12 所示。

图 7- 12：新增项目组信息

新增项目组需输入项目组名称、项目文件的保存目录和描述即可创建项目组。其中项目组名称不能与已存在的项目组名称重复，程序会自动检测名称是否重复。

文件保存目录指在服务器上，保存项目使用的各种文件的目录，如镜像文件、快照等。建议这个目录设置在数字有机体文件系统中。

7.2.2.3 修改项目组

在项目组管理主页面列表中选择一个项目组，点击上部工具栏上的“修改项目组”按钮，弹出修改项目组对话框，如图 7-13 所示。

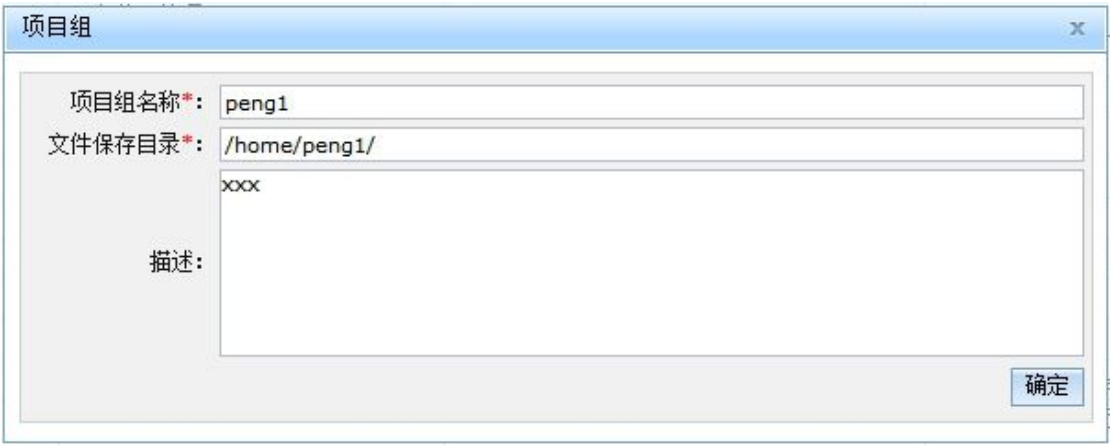


图 7- 13：修改项目组信息

修改项目组操作可修改项目组的名称、项目文件的保存目录和描述信息。

7.2.2.4 删除项目组

在项目组管理主页面列表中选择一个项目组，点击上部工具栏上的“删除”项目组按钮，弹出确认删除对话框，如图 7-14 所示，点击确定即可删除项目组。

图 7- 14：确认删除项目组

如果项目组还有项目，则必须先从项目组中删除这些项目后才能删除该项目组，否则将报错。

7.2.2.5 新增项目到项目组

在项目组管理主页面列表中选择一个项目组，点击上部工具栏上的“新增项目到项目组”按钮，弹出新增项目到项目组对话框，如图 7-15 所示。

图 7- 15：新增项目到项目组

在该页面上“项目名称“下拉框中选择一个项目，在下面的”加入到项目组“下拉框中选择一个项目组，点击确定即可将该项目加入到该项目组。

7.2.2.6 从项目组删除项目

项目组管理主页面列表中选择一个项目组/项目记录，点击上部工具栏上的“从项目组删除项目”按钮，即可将项目从项目组中删除。

7.3项目资源使用情况查看

7.3.1 系统管理员项目资源使用情况查看

系统管理员登陆系统后，点击上部导航栏的“项目管理”进入系统管理员“项目资源使用情况”查看页面，如图 7-16 所示。

图 7- 16：系统管理员项目资源使用情况查看

系统管理员看到的项目资源使用情况查看页面上部为查询输入栏，中部为项目资源使用情况查看列表，底部为项目资源使用情况查看列表分页显示栏。

项目资源使用情况列表显示了项目资源使用信息。

1) 最大虚拟机实例数指项目允许的创建的最大虚拟机实例数；
2) 已建实例数指已经创建的实例数；
3) 最大 VCPU 数指虚拟机允许的最大 VCPU 数；
4) 最大内存数指项目所有虚拟机的内存数的总和；
5) 已使用内存数指项目已经使用的内存数总和；
6) 最大磁盘容量指项目允许的最大磁盘容量；
7) 已使用磁盘容量指项目已经使用的磁盘容量。

7.3.2 项目管理员项目资源使用情况查看

项目管理员登录系统后，点击上部导航栏的“资源使用情况”进入项目资源使用情况查看页面，如图 7-17 所示。

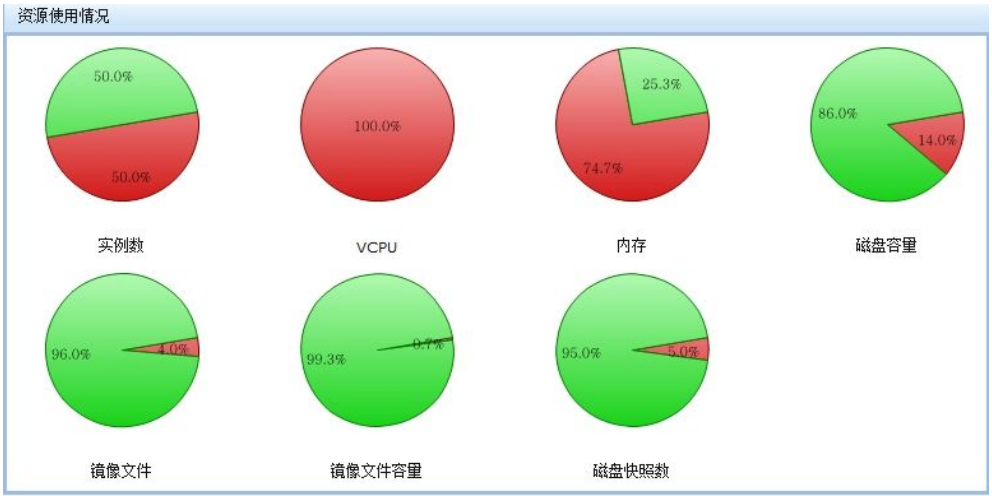


图 7- 17：资源使用情况

页面显示了项目资源的使用情况（使用比例）的饼图，如虚拟机实例数、VCPU 数、内存数、磁盘容量、镜像文件数、镜像文件容量和磁盘快照数。其中红色表示已使用，绿色表示未使用。当鼠标移动到饼图上时，将显示资源使用情况的详细数据。

8. 安全对象管理

8.1 安全对象概述

在数字有机体虚拟机系统中，密码、密钥等信息并不直接设置到访问对象中。例如，在访问 IP-SAN 系统时，除了采用客户端 IP 地址认证外，通常还采用用户/口令认证。为了增强口令存储的安全，将访问口令作为一个独立的对象看待，称作安全对象。每个安全对象的密钥信息都经过加密再存储，而且不能获取安全对象的密钥信息。

在数字有机体虚拟机系统中，安全对象用于保存访问 iSCSI 存储的密码或者证书、卷加密密钥、网络文件系统访问密钥等。这些安全对象最终将传递给宿主机，在宿主机上创建为 libvirt 的安全对象。在配置存储池时，可以将安全对象应用到具体的存储池。

8.2 系统管理员管理安全对象

系统管理员登录系统后，在导航栏中点击“安全对象”管理按钮即可进入安全对象管理的界面，如图 8-1 所示。

图 8- 1：安全对象管理界面

进入该部分的首页是“安全对象信息列表”。这个页面以列表的形式列出系统所有安全对象，分页显示。列表的顶端是自定义查询工具栏。可以按照中文名称、英文名称、项目和用途进行查询。列表的字段为：中文名称、英文名称、用途、所属项目和描述。工具栏的按钮有新增、修改和删除。

点击“新增”按钮时，即弹出新增安全对象面板（图 8-2）。系统管理员输入安全对象的中文名、英文名、用途类型、用途和描述，然后设定其密码。用途类型当前仅有 volume、ceph 和 iscsi 三种类型。volume 类型用于存储卷的加密，即安全对象保存的是卷的加解密密钥。ceph 类型用于访问 ceph 对象存储系统。数字有机体虚拟机系统不支持 ceph，因此不要选择。iscsi 类型用于 iSCSI 存储池的访问。安全对象中保存的是访问 iscsi 系统的密钥或者密码。

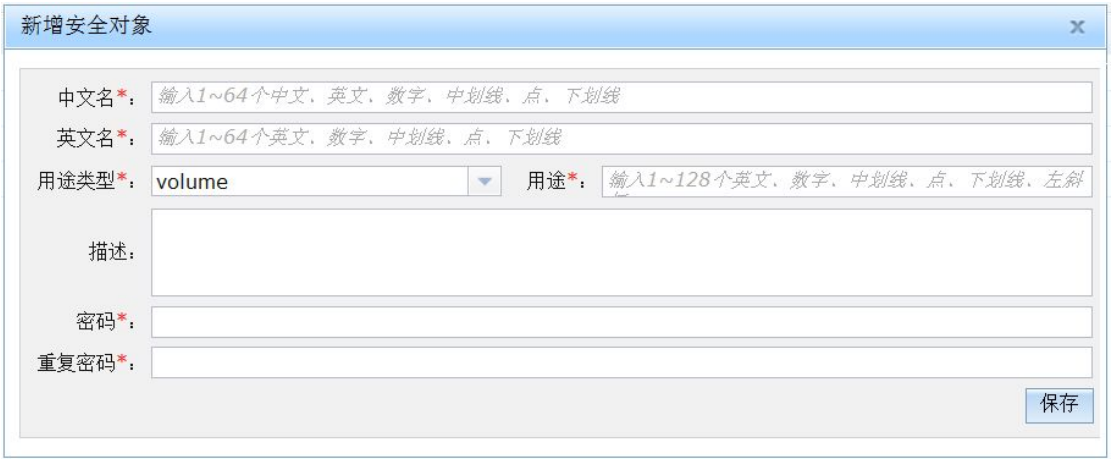


图 8- 2：新增安全对象

点击“修改”按钮时，即弹出修改安全对象面板信息的对话框，如图 8-3 所示。修改安全对象时，系统管理员可以修改其中文名、英文名、用途和描述，以及修改它的值。

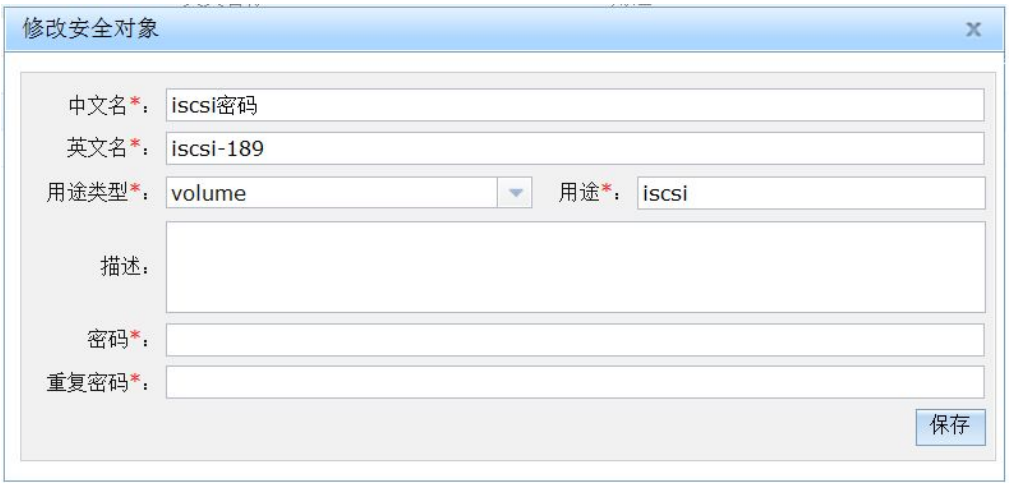


图 8- 3：修改安全对象

删除安全对象时，首先向用户确认，由用户点击确认后做相应操作，并返回操作结果。如果安全对象还在被使用，则删除将失败。

9. 存储虚拟化管理

9.1 存储虚拟化概述

通常，虚拟机并不独占具体的某个物理存储设备，而是使用共享的，或者说虚拟的存储空间。存储虚拟化就是要为虚拟机提供这样的共享的，虚拟的存储空间，但对虚拟机来说，它就像获得了一个独占的存储设备一样。

存储虚拟化实现的思路大致分为两类。一类是以物理设备分区共享为基础的存储虚拟化技术。由 FC-SAN 和 IP-SAN 实现的存储虚拟化方案就属于这类。另外，在单服务器上使用的磁盘卷池和逻辑卷池也属于这一类。这类存储虚拟化技术向虚拟机提供的是固定大小的，物理上存在的逻辑设备。虚拟机模拟器直接访问这个逻辑设备，将其作为虚拟机的磁盘设备。为了统一管理，仍然将这些物理存储设备视作存储池，并把其中的可以独立分配的单元视作卷。如 SAN 存储池提供的 LUN 就是一个卷，而在磁盘卷池中则一个分区就是一个卷。

另一类则是以软件模拟技术为基础的存储虚拟化方案。这类方案依靠虚拟机模拟器中集成的磁盘虚拟软件实现。这些磁盘虚拟软件通常使用文件（一个或者多个）来模拟磁盘设备。为了实现磁盘模拟，文件必然是有格式的。因此又将这些文件称作磁盘镜像文件。不同的磁盘虚拟软件采用不同的文件格式。本系统主要用 qemu 模拟器来模拟磁盘。它支持许多种格式的磁盘镜像文件，如常用的 raw、qcow、qcow2 等。既然模拟软件使用文件来存储磁盘数据，则可以将文件存储在不同的文件系统中。如果存储在宿主机本地文件系统中，则将其称为本地文件目录池。如果存储在网络文件系统中，如 NFS 或者 CIFS，则称其为网络文件系统池，如果存储在数字有机体文件系统中，则称其为数字有机体目录池。

数字有机体虚拟机系统支持多种方式的存储虚拟化方案，而且可以将不同的存储虚拟化方案整合到系统中，以满足不同应用的需求。具体内容请参见 3.2 节“存储虚拟化方案”的描述。

9.2 存储管理概述

数字有机体虚拟机系统支持以下类型的存储池：

1）数字有机体目录池：在数字有机体文件系统目录下创建的存储池，其中的卷可以在任意宿主机上访问，能够被共享，也支持虚拟机迁移。

2）网络文件系统池：包括两种小类型：nfs 和 cifs。现有的 NAS 存储系统都支持 nfs 和 cifs 协议，因此可以基于 NAS 存储系统实现网络文件系统池。也可以单独建立网络文件服务系统。例如利用 Linux 或者 Unix 系统构建的 NFS 文件服务器可以提供 nfs 文件共享，实现 nfs 类型的网络文件系统池。利用 Windows 系统可以部署基于 cifs 的文件共享服务，即可用于实现 cifs 网络文件系统池。通常，需要将网络文件服务挂载到宿主机的某个目录下，然后由管理系统在该目录下访问网络文件服务。其中的卷也可以在任意宿主机上访问，能够被共享，也支持虚拟机迁移。

3）iSCSI 池：这是为了支持 IP-SAN 这样的存储系统实现的。其中的卷也可以在任意宿主机上访问，也支持虚拟机迁移。

4）宿主机文件目录池：在宿主机本地文件系统内构建目录池。这种池是绑定宿主机的，在其他宿主机上无法访问。使用这种池内的卷的虚拟机也相应地绑定在该台宿主机上。

5）逻辑卷池、磁盘卷池：这两个池都是使用宿主机上已有的磁盘设备来构建。前者使用的是磁盘分区，后者使用的是磁盘设备。而且这两个池都是绑定宿主机的。

卷是指存储在存储池中的磁盘镜像文件，或者物理上的卷设备。要注意的是：数字有机体目录池、网络文件系统池、宿主机文件目录池是支持创建、删除、扩展、转存卷为镜像、创建卷快照操作的。iSCSI 池、逻辑卷池、磁盘卷池则只能查看卷和修改卷名称和描述，其他操作都不能使用。

9.3 系统管理员管理存储池和卷

只能由系统管理员来管理系统的存储池，然后将存储池分配给项目使用。存储池管理的内部网页结构如图 9-1 所示。

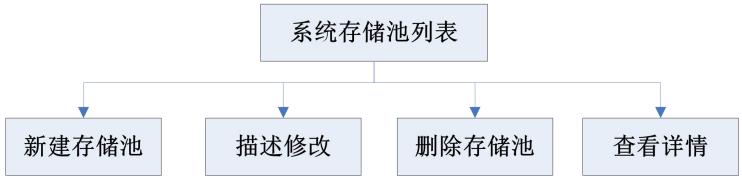


图 9- 1：存储池管理的内部结构

9.3.1 系统存储池信息

存储池管理的首页面是“系统存储池信息”，如图 9-2 所示。这个页面以列表的形式列出系统的所有存储池。系统管理员可以根据存储池名称、逻辑容量大小、类型、创建时间、是否绑定宿主机、绑定宿主机名称等条件过滤数据。数据结果分页显示。列表顶端有五个功能按钮：存储卷信息、新增、修改描述、删除和查看详情。

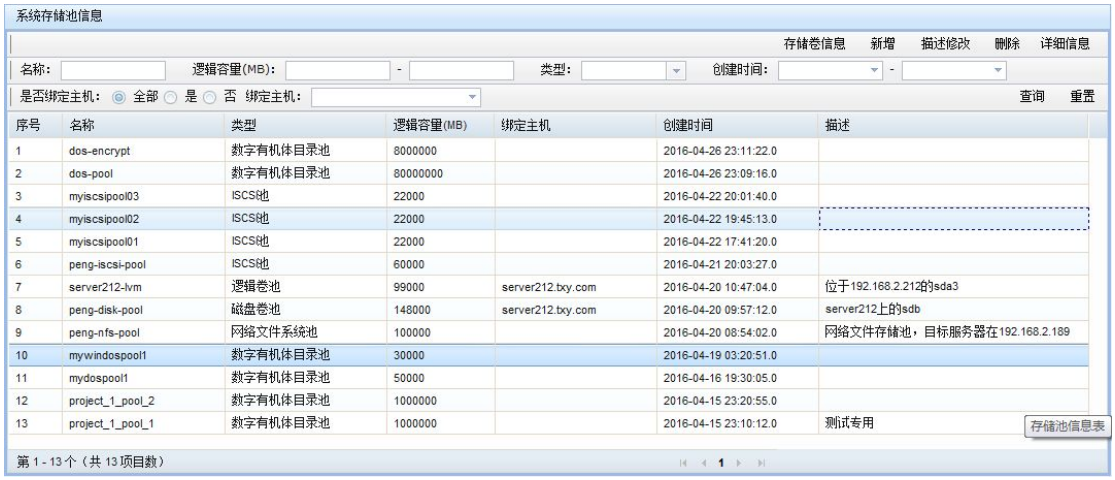


图 9- 2：存储池信息页面

9.3.2 新增存储池

本系统支持了六种存储池：数字有机体目录池、网络文件系统池、iSCSI 池、宿主机文件目录池、逻辑卷池、磁盘卷池。由于逻辑卷池和磁盘卷池的配置相同，因此下面分五种情况来详细描述存储池的创建。

9.3.2.1 数字有机体目录池

数字有机体目录池构建在数字有机体文件系统中。数字有机体系统中已经集成了数字有机体文件系统。请参考《数字有机体工作平台及抗毁容灾系统用户手册》部署数字有机体文件系统。

数字有机体文件系统利用每台服务器共享出的存储空间构建虚拟的存储系统，并在这个存储系统中建立共享的文件系统。如果每台服务器（宿主机）都是数字有机体文件系统中的一个节点，则每台服务器都能安装数字有机体文件系统到本地文件系统中，并且看到同样的文件系统内容。因此，在数字有机体文件系统中建立的存储池可以在每台服务器上访问到。如果虚拟机使用数字有机体目录池中的卷，则它可以不受卷所在的位置约束，因为卷本身就是可以在每台服务器上访问到得。

数字有机体目录池使用文件来模拟磁盘。通常每个文件就是一个卷。这些卷存储在指定的数字有机体目录下。项目管理员可以在存储池中创建新的卷，扩展卷的容量、为卷制作快照和删除卷。删除卷时将同时删除文件。

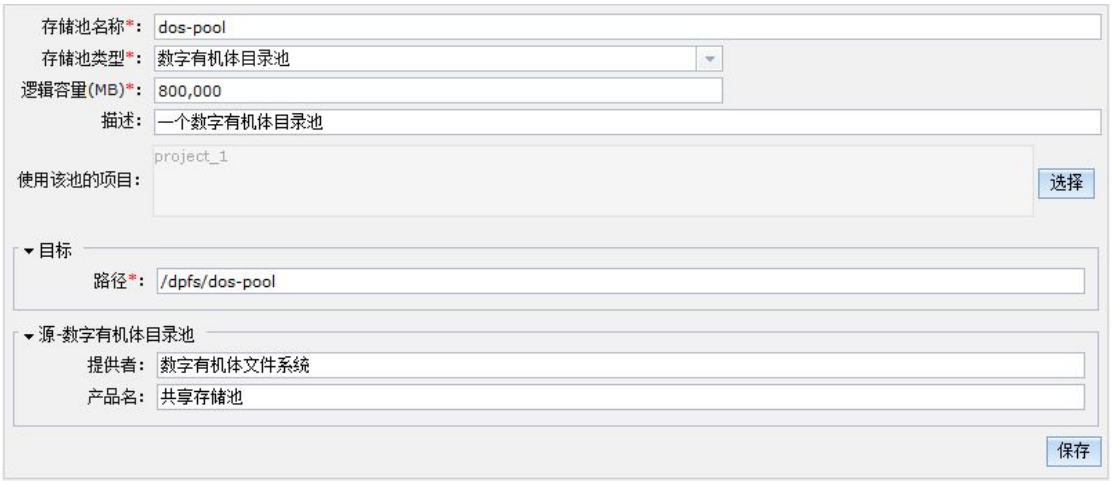


图 9- 3：数字有机体目录池的配置内容

新增数字有机体目录池的配置内容如图 9-3 所示。各项的设置和含义如下：

1) 存储池名称：存储池在系统中的唯一名称，只允许字母、数字、中划线、下划线和点。
2) 存储池类型：存储池的类型，此处的值为数字有机体目录池。
3) 逻辑容量：存储池下面所有卷的容量总和的最大值。在创建卷时如果使用的容量加上已有卷的容量和超过这个值，那么卷就会创建失败。
4) 描述：对当前存储池的详细说明，仅用于方便用户对此存储池的了解。
5) 使用该池的项目：设置能使用该池的项目，可以在项目管理中修改。
6) 源-路径：用于指定当前存储池将映射到本地文件系统名字空间的位置。对于数字

有机体目录池，它应是卷将要创建的目录（卷将创建在这个目录下），且它必须是在数字有机体文件系统目录下。

7） 源-提供者：存储设备的提供者，主要用于查看设备时的显示信息。

8） 源-产品名：存储设备的产品名，主要用于查看设备时的显示信息。

9.3.2.2 网络文件系统池

网络文件系统池构建在网络文件系统中。网络文件系统可以采用不同的方式实现。本系统支持基于 nfs 或者 cifs 协议的网络文件系统。

在部署网络文件系统池前，需要先配置好网络文件系统的服务。如果使用某台服务器提供 nfs 服务，则应该在该服务器上配置输出目录，访问权限等。如果使用 windows 服务器提供 cifs 服务，则要在服务器上配置好 windows 的文件共享。如果使用 NAS 存储系统，也需要在 NAS 存储系统上完成相应配置。并且需要获得访问服务需要的参数，包括服务提供者的 IP 地址，服务端口，共享目录的 URL 和访问用户/密码等。

现以一个示例来说明在一台 server212 的服务器上创建 nfs 池和 nfs 卷的过程。

服务器需要开启 nfs 服务（由于 DOS 服务器默认开启 nfs 服务，开启 nfs 服务过程略）。然后在命令行输入“vi /etc/exports”编辑/etc/exports 文件，在文件尾添加“/raid/disk-pool 192.168.2.0/24(rw)”，然后保存，并执行命令“exportfs -fr”，如图 9-4 所示。

```
# /etc/exports: the access control list for filesystems which may be exported
#               to NFS clients.  See exports(5).
#
# Example for NFSv2 and NFSv3:
/srv/nfsboot       *(rw,sync,no_subtree_check)
#
# Example for NFSv4:
# /srv/nfs4        gss/krb5i(rw,sync,fsid=0,crossmnt,no_subtree_check)
# /srv/nfs4/homes  gss/krb5i(rw,sync,no_subtree_check)
#
/raid/disk-pool    192.168.2.0/24(rw)
```

图 9- 4：编辑/etc/exports 文件配置服务器输出目录示例

其中“/raid/disk-pool”是服务器的输出目录，“192.168.2.0/24”是指 nfs 服务的网段（nfs 客户端 IP 地址的范围），“rw”是指可目录可读写，同时默认的权限是 root_squash 和 no_all_squash。在命令行输入“cat /var/lib/nfs/etab”查看输出目录的配置信息，如图 9-5 所示。

```
root@server212:/raid/disk-pool# cat /var/lib/nfs/etab
/raid/disk-pool 192.168.2.0/24(rw,sync,wdelay,hide,nocrossmnt
,secure,root_squash,no_all_squash,no_subtree_check,secure_loc
ks,acl,anonuid=65534,anongid=65534,sec=sys,rw,root_squash,no_
all_squash)
/srv/nfsboot    *(rw,sync,wdelay,hide,nocrossmnt,secure,root_
squash,no_all_squash,no_subtree_check,secure_locks,acl,anonui
d=65534,anongid=65534,sec=sys,rw,root_squash,no_all_squash)
/raid/data      *(rw,sync,no_wdelay,nohide,nocrossmnt,secure,
no_root_squash,no_all_squash,no_subtree_check,secure_locks,ac
l,anonuid=-2,anongid=-2,sec=sys,rw,no_root_squash,no_all_squa
sh)
```

图 9- 55：查看服务器输出目录配置信息示例

在 nfs 客户端主机上创建目录/media 作为 nfs 服务器输出目录/raid/disk-pool 映射到本地文件系统的位置。

用系统管理员账号登陆数字有机体虚拟机管理系统，进入存储池管理页面，点击新建按钮，进入新建存储池页面，创建 nfs 存储池各项配置输入如图 9-6 所示。

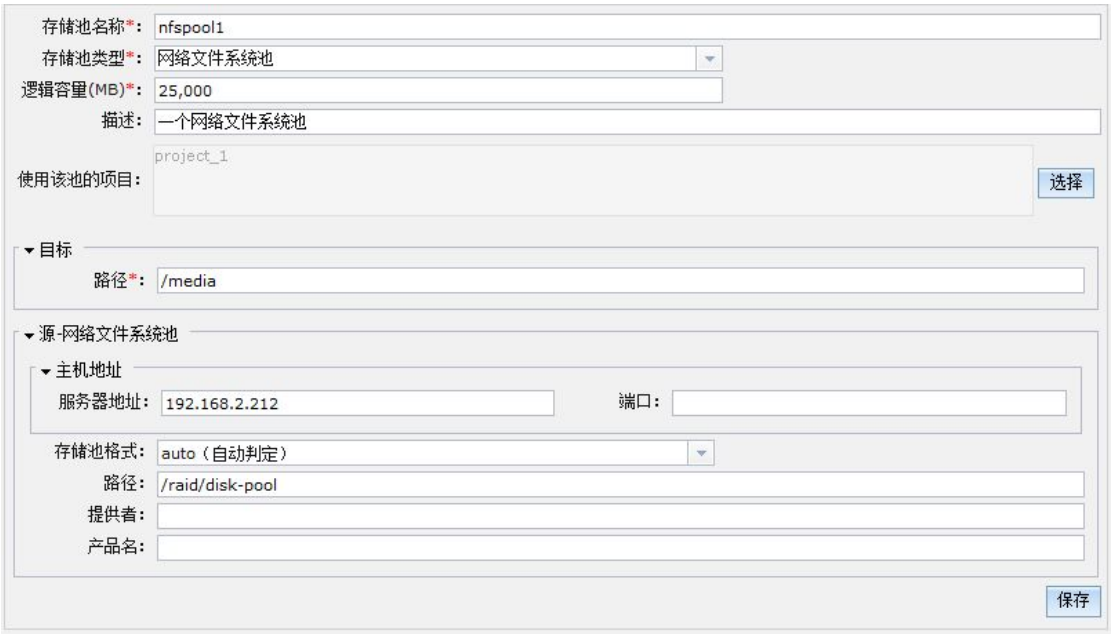


图 9- 6：网络文件系统池的配置示例

网络文件系统池的配置内容如图 9- 所示，各配置项的含义如下：

1) 存储池名称：存储池在系统中的唯一名称，只允许字母、数字、中划线、下划线和点。此处值为 nfspool1。
2) 存储池类型：存储池的类型，此处值为网络文件系统池。
3) 逻辑容量：存储池下面所有卷的容量总和的最大值。在创建卷时如果使用的容量加上已有卷的容量超过这个值，那么卷就会创建失败。此处值为 25000MB。
4) 描述：对当前存储池的详细说明，仅用于方便用户对此存储池的了解。
5) 使用该池的项目：设置能使用该池的项目，可以在项目管理中修改。
6) 目标-路径：用于指定当前存储池将映射到本地文件系统名字空间的位置。对于网络文件系统池，它应是卷将要创建的目录（卷将创建在这个目录下）。此处值为/media。
7) 源-主机地址-服务器地地址：网络文件系统的服务器地址，可以是服务器的主机名或者 IP 地址。此处值为 192.168.2.212。
8) 源-主机地址-端口：服务器的访问端口。可以不配置，由系统使用缺省值。
9) 源-存储池格式：网络文件系统的协议类型。现支持以下类型：auto（自动判定）、nfs、glusterfs 和 cifs。nfs：是 FreeBSD、Linux 等支持的文件系统中的一种，它允许网络中的计算机之间通过 TCP/IP 网络共享资源。glusterfs：它是一个开源的分布式文件系统，具有强大的横向扩展能力，通过扩展能够支持数 PB 存储容量和处理

数千客户端，它借助 TCP/IP 或 InfiniBand ADMA 网络将物理分布的存储资源集中在一起，使用单一全局命名空间来管理数据。cifs：使程序可以访问远程 Internet 计算机上的文件并要求此计算机提供服务，它使用客户/服务器模式。此处值为 auto。

10) 源-路径：存储池的远地访问目录名。对 NFS 来说，就是服务器上的输出目录名，对 cifs 来说，就是共享目录名。此处值为/raid/disk-pool。
11) 源-提供者：存储设备的提供者，主要用于查看设备时的显示信息。可不填。
12) 源-产品名：存储设备的产品名，主要用于查看设备时的显示信息。可不填。

新建 nfs 存储池后，用项目管理员账号登陆数字有机体管理系统，进入存储卷管理页面，在 nfs 存储池 nfspool1 中创建 nfs 存储池卷 nfsvol1（逻辑容量 10000MB），具体创建方法请参考 9.4.3 小节。后台程序在 nfs 客户机目录/media（映射到服务器输出目录/raid/disk-pool）创建了卷 nfsvol1，由于服务器输出目录/raid/disk-pool 的权限前面设置的为权限默认值，即 root_squash 和 no_all_squash，故卷 nfsvol1 的所有者为 nobody，如图 9-7 所示。

```
root@server212:/raid/disk-pool# l
总用量 196
-rw------- 1 nobody nogroup 197120 4月    5 14:40 nfsvol1
```

图 9- 7：创建的 nfs 卷的权限信息示例

该卷文件的访问权限为-rw- r-- r--，需在命令行输入“chmod 777 nfsvol1”命令将该卷文件的访问权限更改为-rwx rwx rwx，否则后台程序回报“/media/nfsvol1:Permission denied”错误。

在虚拟机管理系统创建基于 nfs 存储池的虚拟机实例后，创建虚拟机并运行虚拟机，运行虚拟机选择宿主机后，后台程序自动在宿主机（服务器，作为 nfs 客户端）创建服务器本地目录/media 到 nfs 服务器本地目录/raid/disk-pool 的映射。例如宿主机选择 server185，nfs 服务器选择 server212，在 server185 上命令行输入 mount 命令，将看到如图 9-8 所示的目录映射信息。

```
192.168.2.212:/raid/disk-pool on /media type nfs4 (rw,relatim
e,vers=4,rsize=262144,wsize=262144,namlen=255,hard,proto=tcp,
port=0,timeo=600,retrans=2,sec=sys,clientaddr=192.168.2.185,m
inorversion=0,local_lock=none,addr=192.168.2.212)
```

图 9- 8：nfs 目录映射信息示例

9.3.2.3 iSCSI 池

iSCSI 池构建在通过 iSCSI 协议可以访问的存储系统上。IP-SAN 系统都支持通过 iSCSI 协议来访问，因此 iSCSI 池通常建立在 IP-SAN 系统中。当然，某些系统也支持通过 iSCSI 协议访问虚拟的存储设备，不过这并不是 iSCSI 常用的方式。

在配置 iSCSI 池前，需要先配置好 IP-SAN 系统。部署 IP-SAN 系统的方案可以参见第三章的描述。在部署和配置 IP-SAN 系统时，记下访问存储系统需要的地址、端口、用户名和密码。

iSCSI 池的存储服务由网络上的服务系统提供。因此，只要宿主机能够访问服务系统，宿主机即可建立相同的存储池。所以，iSCSI 池并不绑定具体的宿主机。但是，需要确保系统中的每台服务器都能访问存储服务系统。

iSCSI 池中的卷是 IP-SAN 系统中分配的逻辑单元，即 LUN。通常，在一个 target 下可以划分出许多个 LUN，每个 LUN 作为一个独立的卷给虚拟机使用。系统还不支持直接控制 IP-SAN 系统，因此无法在 iSCSI 池下创建和修改卷，也不能为其中的卷创建快照。需要在 IP-SAN 系统中分配好 LUN，然后通过管理界面中的存储池刷新来获得卷的信息（每个 LUN 对应一个卷）。

可以用服务器模拟 iSCSI 设备。在服务端（tgt 端）服务为 tgt，服务进程为 tgtd。可以使用 sevice tgt start 来启动，也可以直接执行/etc/init.d/tgt start 来启动。客户端的服务为 open-iscsi，可以用 service open-iscsi start 命令来启动它。然后就可以用 iscsiadm 命令来查询和登录 tartget。

下面分别说明服务端和客户端的配置。

（1）服务端配置

先配置一个 target，名称可根据自己的需要选择，id 从 1 开始。例如：

tgtadm --lld iscsi --op new --mode target --tid 1 --targetname qyj-target

上面的命令创建了一个 target，id 为 1，名称为 qyj-target。

然后可以为 tartget 增加 LUN。一个 LUN 可以是模拟镜像、真实设备或者 passthrough 设备等。这里先看一个模拟镜像的使用：

1）创建一个模拟镜像文件：

tgtimg --op new --device-type disk --type disk --size 10000 --file /home/vmbgdisk.img

这里创建一个 10G 的磁盘镜像文件，保存在/home/vmbgdisk.img 中。

2）将镜像文件指定为一个 LUN

tgtadm --lld iscsi --mode logicalunit --op new --tid 1 --lun 1 -b /home/vmbgdisk.img --blocksize=4096

这里用-b 选项指定镜像文件和块大小。如果使用物理设备作为 LUN 则可以用：-b /dev/sda。

3）bind target

只有 bind 了的 tartget 才能被 initiator 访问。

tgtadm --lld iscsi --op bind --mode target --tid 1 -I ALL

-I 指定允许访问者，这里用 ALL 表示所有的。如果不希望按照客户地址来限定，可以采用账号/口令方式，命令形如：

tgtadm --lld iscsi --mode account --op new --user qgh --password 123456

tgtadm --lld iscsi --mode account --op bind --tid 1 --user qgh（绑定账号 qgh 到 target 1）

tgtadm --lld iscsi --mode account --op unbind --tid 1 --user qgh（若想解除用户绑定 target，可以用该命令解除账号 qgh 与 target 1 的绑定）

可以用“tgtadm --lld iscsi --op show --mode target”命令查看创建的 tartget 的信息。如图 9-9 所示。

```
root@server189:/home/dos# tgtadm --lld iscsi --op show --mode target
Target 1: qyj-target
    System information:
        Driver: iscsi
        State: ready
    I_T nexus information:
        LUN: 2
            Type: disk
            SCSI ID: IET     00010002
            SCSI SN: beaf12
            Size: 10486 MB, Block size: 512
            Online: Yes
            Removable media: No
            Prevent removal: No
            Readonly: No
            SWP: No
            Thin-provisioning: No
            Backing store type: rdwr
            Backing store path: /home/vmbgdisk.img
            Backing store flags:
    Account information:
        qyj
        qgh
        aaa
        bbb
    ACL information:
        ALL
```

图 9- 9：创建的 target 信息示例

4）可以用如下的命令将配置保存到配置文件中。否则重启 tgtd 后将丢失这些配置。

tgt-admin --dump >/etc/tgt/conf.d/qyj-target.conf

这样，重新启动是将再读取这些配置文件并自动完成配置。

可使用 cat /etc/tgt/conf.d/qyj-target.conf 查看保存的配置文件信息。如图 9-10 所示。

```
root@server189:/home/dos# cat /etc/tgt/conf.d/qyj-target.conf
default-driver iscsi

<target qyj-target>
        backing-store /home/bgdisk.img
        backing-store /home/vmbgdisk.img
        incominguser qyj PLEASE_CORRECT_THE_PASSWORD
        incominguser qgh PLEASE_CORRECT_THE_PASSWORD
        incominguser aaa PLEASE_CORRECT_THE_PASSWORD
        incominguser bbb PLEASE_CORRECT_THE_PASSWORD
</target>
```

图 9- 10：创建的 target 配置信息示例

（2）客户端配置（initiator 端）

客户端的服务为 open-iscsi，可以用 service open-iscsi start 命令来启动它。然后就可以用 iscsiadm 命令来查询和登录 tartget。

查询 tartget 的命令形如：

iscsiadm --mode discoverydb --type sendtargets --portal 192.168.2.189 --discover

“--portal 后指定 target 服务器的地址，如果使用默认端口可以不指定，否则使用“192.168.2.189:3260”这样的形式。该命令的执行结果如图 9-11 所示。

```
root@server185:/dev/disk/by-path# iscsiadm --mode discoverydb -
-type sendtargets --portal 192.168.2.189 --discover
192.168.2.189:3260,1 qyj-target
```

图 9- 11：iSCSI 客户端查询服务端创建的 target 信息示例

上图显示 target 服务端 IP:端口为 192.168.2.189:3260，target 名称为 qyj-target，target ID 为 1。

虚拟机管理系统创建 iSCSI 存储池的示例如图 9-12 所示。

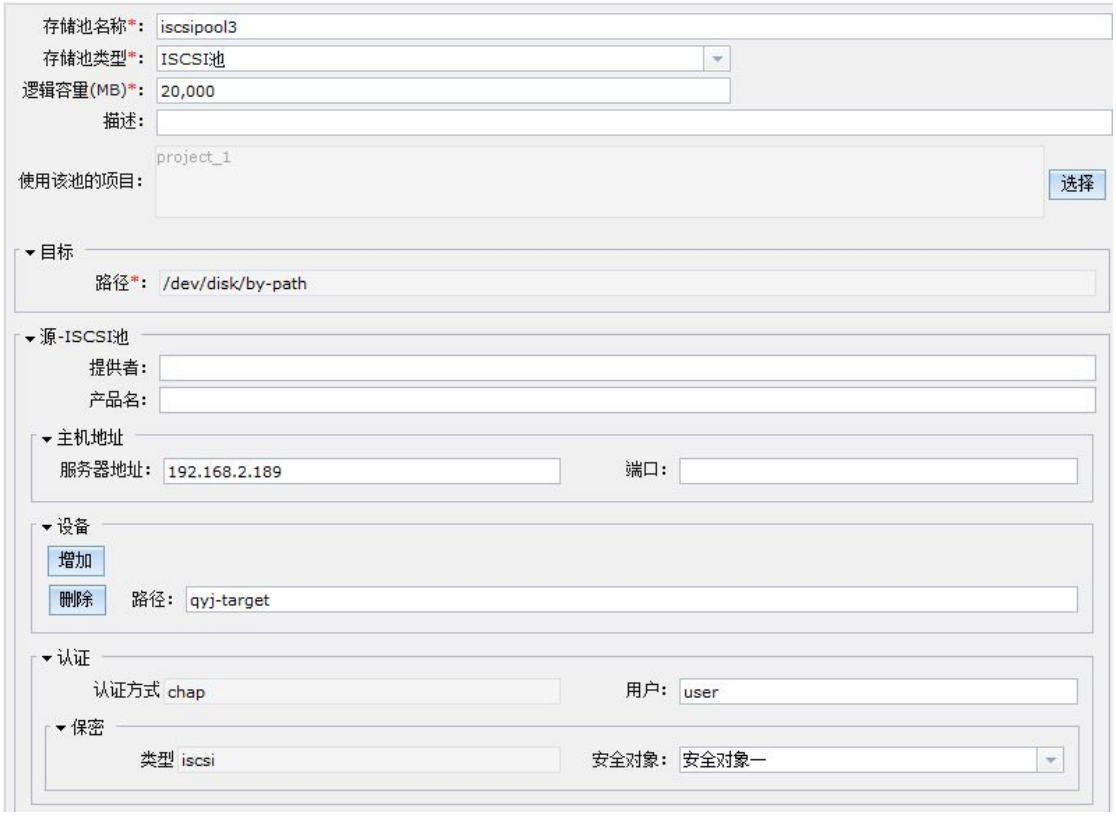


图 9- 12：iSCSI 存储池的配置示例

iSCSI 池的配置内容如图 9-12 所示。各个配置项的含义如下：

1) 存储池名称：存储池在系统中的唯一名称，只允许字母、数字、中划线、下划线和点。
2) 存储池类型：存储池的类型，此处值为 iSCSI 池。
3) 逻辑容量：对 iSCSI 池来说没有意义，可以设置为可用容量。
4) 描述：对当前存储池的详细说明，仅用于方便用户对此存储池的了解。
5) 使用该池的项目：设置能使用该池的项目，可以在项目管理中修改。
6) 目标-路径：用于指定当前存储池将映射到本地文件系统名字空间的位置。对于 iSCSI 池，它应是存放设备文件的目录。由于重启系统并不保证设备文件是恒定的，因此此处使用固定值/dev/disk/by-path。
7) 源-提供者：存储设备的提供者，主要用于查看设备时的显示信息。
8) 源-产品名：存储设备的产品名，主要用于查看设备时的显示信息。
9) 源-主机地址-服务器地址：后台存储的远地服务器地址，可以是服务器的主机名或者 IP 地址。
10) 源-主机地址-端口：服务器的访问端口。
11) 源-设备-路径：对 iSCSI 池来说，这里是要访问的 tartget 的名称。

12）源-认证-认证方式：访问远地服务器的认证方式，此处固定值 chap。

13）源-认证-用户：标志认证中使用的用户。如果不需要认证可不填。

14）源-认证-保密-类型：固定值 iscsi。

15）源-认证-保密-安全对象：选择要使用的已经配置好的安全对象，其中保存了认证使用的密钥。如果不需要认证可不填。

上图的输入的用户应为上面 target 服务端创建并绑定到 target 1 的用户账号，为 qgh。安全对象应选择 iSCSI 安全对象，注意安全对象的值应与账号 qgh 的密码相同。创建 iSCSI 存储池后，将该存储池添加到项目中。在创建虚拟机实例时，在配置磁盘时选择存储池时选择 iscsipool3，卷选择卷名称为 unit:0:0:2 的卷（即对应于上面在 target 服务端配置时创建的模拟镜像文件/home/vmbgdisk.img）。运行虚拟机是选择宿主机时，例如选择 server185，则系统自动用命令“iscsiadm --mode node --targetname qyj-target --portal 192.168.2.189:3260 --login”登陆到 target 服务端（server185 作为客户端），用户验证使用的用户名和密码为上面页面配置是设定的用户名和安全对象的值。在 server185 上用 fdisk –l 命令可以看到一个新增加的模拟设备/dev/sdc，如图 9-13 所示。

```
Device     Boot  Start      End  Sectors  Size Id Type
/dev/sdc1  *      2048   206847   204800  100M  7 HPFS/NTFS/exFAT
/dev/sdc2       206848 20477951 20271104  9.7G  7 HPFS/NTFS/exFAT
```

图 9- 13：iSCSI 服务端创建的模拟设备信息示例

在 server185 上的目录/dev/disk/by-path 可以看到上面在 target 服务端配置时创建的模拟镜像文件/home/vmbgdisk.img（lun 2），如图 9-14 所示。

```
root@server185:/dev/disk/by-path# ls
ip-192.168.2.189:3260-iscsi-qyj-target-lun-1
ip-192.168.2.189:3260-iscsi-qyj-target-lun-2
```

图 9- 16：iSCSI 创建的 lun 信息示例

9.3.2.4 宿主机文件目录池

宿主机文件目录池建立在宿主机的文件系统中，即在宿主机的文件系统的某个目录下建立存储池。和数字有机体目录池、网络文件系统池一样，池中的卷用文件来模拟，因此需要 qemu 的支持。通常，这类存储池和具体的宿主机绑定，因为指定的目录是这台宿主机上的，其他宿主机上无法访问到。

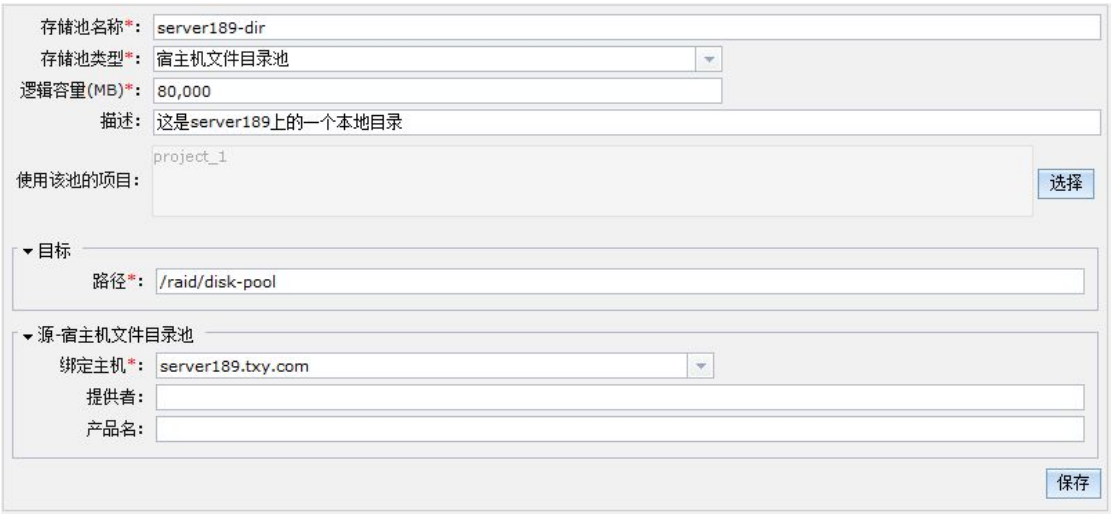


图 9- 15：宿主机文件目录池的配置示例

宿主机文件目录池的配置内容如图 9-所示。每个配置项的含义如下：

1) 存储池名称：存储池在系统中的唯一名称，只允许字母、数字、中划线、下划线和点。
2) 存储池类型：存储池的类型，此处值宿主机文件目录池。
3) 逻辑容量：存储池下面所有卷的容量总和的最大值。在创建卷时如果使用的容量加上已有卷的容量超过这个值，那么卷就会创建失败。
4) 描述：对当前存储池的详细说明，仅用于方便用户对此存储池的了解。
5) 使用该池的项目：设置能使用该池的项目，可以在项目管理中修改。
6) 目标-路径：用于指定当前存储池将映射到本地文件系统名字空间的位置。对于宿主机文件目录池，它应是卷将要创建的目录（卷将创建在这个目录下）。
7) 源-绑定主机：指定当前存储池需要绑定的宿主机。
8) 源-提供者：存储设备的提供者，主要用于查看设备时的显示信息。
9) 源-产品名：存储设备的产品名，主要用于查看设备时的显示信息。

9.3.2.5 磁盘卷池

磁盘卷池使用宿主机（就是服务器）上的磁盘设备构建。磁盘设备可以是不同类型的。最简单的，可以使用宿主机上的某个大容量磁盘作为磁盘卷池的构建设备。如果宿主机内置有磁盘阵列卡，则可以将多块磁盘连接到磁盘阵列卡上，由磁盘阵列卡将其整合为一个更大容量的，更好性能的磁盘设备。从某种意义上讲它也是一个虚拟的、逻辑上的磁盘，因此在阵列配置中它被称为逻辑磁盘，但是对操作系统来说，它就是一个标准的磁盘设备。阵列系统通常支持 raid0、raid1、raid5 和 raid6 等。建议采用 raid5 或者 raid6 来整合多块磁盘。

磁盘卷池的磁盘设备也可以是宿主机外接的磁盘阵列。内置的磁盘阵列难以部署几十块硬盘，而磁盘阵列却可以轻松的做到。因此，在需要大容量存储时，可以采用外接磁盘阵列的方式。在部署这类磁盘卷池前，需要先配置好磁盘阵列，并为宿主机构建逻辑设备。许多

磁盘阵列都限制单个逻辑设备的大小（如 2TB），因此宿主机上可能出现多个磁盘设备（存储阵列的每个逻辑设备对应宿主机上的一个磁盘设备）。

在数字有机体系统下，磁盘设备以/dev/sdx 或者/dev/hdx 的方式命名。x 是单个字母，从“a”开始，然后是“b”、“c”……。每个磁盘卷池只使用一个磁盘设备。如果要将多个磁盘设备整合为一个存储池，需要使用其他类型，如逻辑卷池或者目录池。

在虚拟机管理系统中，可以使用磁盘设备建立一个磁盘卷池。磁盘的分区作为磁盘卷池的卷，即一个磁盘分区对应一个卷，虚拟机操作系统就安装在卷中。虚拟机管理系统不支持在页面创建磁盘卷池的卷，磁盘卷只能使用磁盘上已经分好的分区，通过虚拟机管理界面中的存储池刷新来获得卷的信息。注意：在虚拟机实例配置中，选择磁盘卷池的卷时，不能选则服务器操作系统安装的分区。因为磁盘卷池把整个分区作为卷来使用，因此若选择的分区上有数据的话，原数据将被创建的卷覆盖。

例如现在使用 server212 上的/dev/sdb 建立一个磁盘池。新增池的设置如图 9-所示，其中各个配置项的含义如表 9-1 所示。

注意：在虚拟机管理系统的虚拟机实例配置页面的启动故障策略选择 NULL，因为启动故障策略的其他选项只对文件类型的卷才正确，磁盘卷池的卷使用磁盘的一个分区，对应与一个磁盘分区，而不是文件类型的卷（即卷是以文件系统的文件的形式存在）。

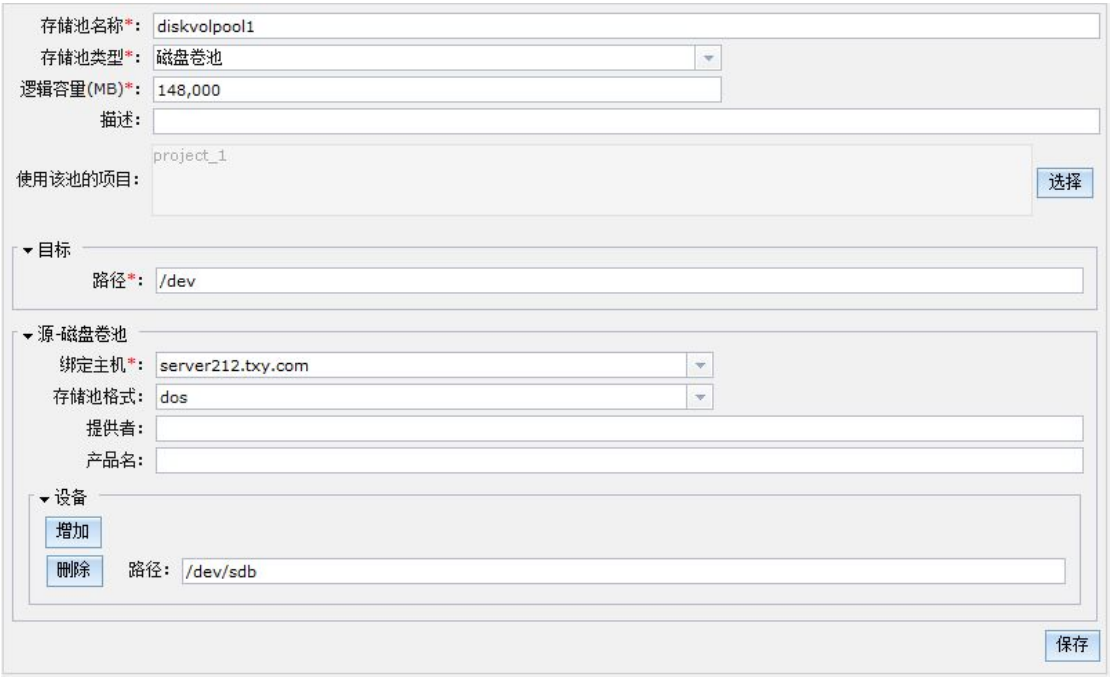


图 9- 16：磁盘卷池的配置示例

表 9-1：磁盘卷池配置项的说明

名称	值	含义
存储池名称	diskvolpool1	存储池的英文名称。
存储池类型	磁盘卷池	
逻辑容量	148000	存储池的总容量大小，需要根据实际情况填写。

目标-路径	/dev	磁盘设备所在的目录，对数字有机体系统来说，总是/dev
源-磁盘卷池-绑定主机	Server212.txy.com	这是选择的，即磁盘所在的服务器。
存储池格式	dos	磁盘分区表的格式，常用的是 dos 和 gpt。如果要使用 2TB 以上的分区，建议使用 gpt。
提供者	Server212	这是一个辅助信息，可选择填写
产品名	磁盘	这是一个辅助信息，可选择填写
设备-路径	/dev/sdb	使用的磁盘的设备路径名。当前只使用第二个磁盘设备，第一个磁盘设备安装操作系统。

9.3.2.6 逻辑卷池

如果宿主机上有多块磁盘，则可以使用逻辑卷来构建存储池。数字有机体系统支持 LVM，可以使用多块磁盘来构建满足各种需求的逻辑卷，如镜像、条带化等。如何使用 LVM 可以参考 LVM 官方手册。

在配置逻辑卷之前，需要先对每块磁盘进行分区。如果要将整块磁盘用作构造 LVM 卷组，则可以只创建一个分区，这个分区使用整个磁盘空间。一个逻辑卷池可以使用多个磁盘分区。

逻辑卷池实际上使用的是一个 LVM 卷组。它使用系统的 vgcreate 命令创建 LVM 卷组，并使用 lvdisplay 命令获得卷组中卷的信息，LVM 卷组中的每个卷对应一个存储池卷。系统还不支持在逻辑卷池中创建卷等。系统管理员需要使用 lvcreate 命令来管理 LVM 卷组中的卷，然后在管理系统中刷新存储池的卷信息，从而获得可用的存储卷。

由于 libvirt 接口的限制，程序不支持自动调用 vgcreate 命令和 lvcreate 命令，只能系统管理员通过命令行手动分别调用 vgcreate 命令和 lvcreate 命令创建 LVM 卷组和 LVM 卷组中的卷。

手动调用 vgcreate 命令创建 LVM 卷组的示例如图 9-17 所示。

```
root@server212:/home/dos# vgcreate server212-lvm /dev/sda3
  Physical volume "/dev/sda3" successfully created
  Volume group "server212-lvm" successfully created
```

图 9-17 创建逻辑卷组示例

Vgcreate 后的 server212-lvm 为 LVM 卷组的名称，也是逻辑卷池的名称，/dev/sda3 是服务器上的 a 磁盘上的第 3 分区，作为所创建的 LVM 卷组的物理设备。注意，由于创建 LVM 卷组使用整个第 3 分区，如果原来的第 3 分区有数据的话，将会丢失，即数据会被在逻辑卷上存储的虚拟机数据所覆盖。一个逻辑卷作为虚拟机的逻辑（虚拟）磁盘。

LVM 卷组创建后，可手动调用 vgdisplay 命令显示 LVM 卷组的信息，如图 9-18 所示。

```
root@server212:/home/dos# vgdisplay server212-lvm
  --- Volume group ---
  VG Name               server212-lvm
  System ID
  Format                lvm2
  Metadata Areas        1
  Metadata Sequence No  1
  VG Access             read/write
  VG Status             resizable
  MAX LV                0
  Cur LV                0
  Open LV               0
  Max PV                0
  Cur PV                1
  Act PV                1
  VG Size               98.77 GiB
  PE Size               4.00 MiB
  Total PE              25285
  Alloc PE / Size       0 / 0
  Free  PE / Size       25285 / 98.77 GiB
  VG UUID               nGDKIr-2Ek2-B0HD-Z0dp-Djzr-NHNB-3iheLE
```

图 9-18 逻辑卷组信息示例

其中 VG Size 的值为 98.77GiB，即是 LVM 卷组的大小，也是/dev/sda3 分区的大小，LVM 卷组在创建时会使用整个分区的磁盘空间。

LVM 卷组创建后，就可以手动调用 lvcreate 命令在卷组中创建逻辑卷，示例如图 9-19 所示。

```
root@server212:/home/dos# lvcreate -L 20G server212-lvm
  Logical volume "lvol0" created
```

图 9-19 创建逻辑卷示例

其中 20G 为创建的逻辑卷的大小，lvol0 为创建卷的名称，由程序自动生成。

逻辑卷创建后可用手动调用 lvdisplay 名称显示创建的卷的信息，如图 9-20 所示。

```
root@server212:/home/dos# lvdisplay server212-lvm
  --- Logical volume ---
  LV Path                /dev/server212-lvm/lvol0
  LV Name                lvol0
  VG Name                server212-lvm
  LV UUID                n8XHC2-SghC-IVK5-mZc9-RIc9-7GRt-EDKi4P
  LV Write Access        read/write
  LV Creation host, time server212, 2016-04-01 10:16:36 +0000
  LV Status              available
  # open                 0
  LV Size                20.00 GiB
  Current LE             5120
  Segments               1
  Allocation             inherit
  Read ahead sectors     auto
  - currently set to     256
  Block device           254:0
```

图 9-20 逻辑卷信息示例

系统管理员通过命令行手动调用命令创建逻辑卷组和逻辑卷后，需要在逻辑卷池创建页面中，输入创建按逻辑卷组（逻辑卷池）的信息创建逻辑卷组。逻辑卷池的配置示例如图 9-21 所示。其中的配置项如表 9-2 所述。

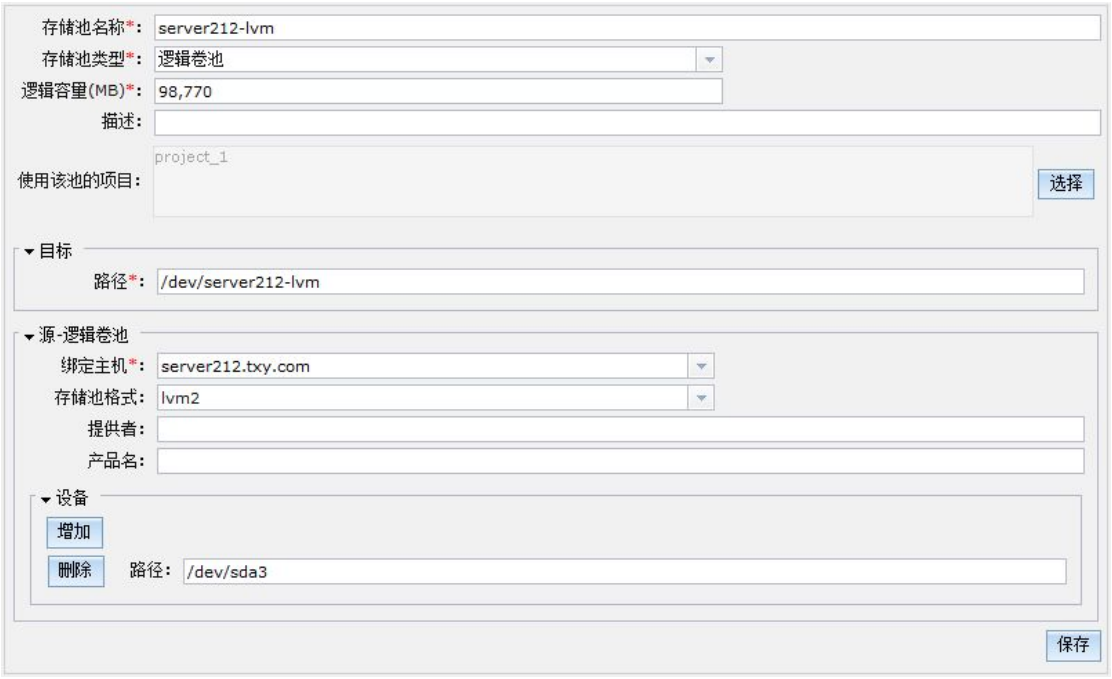


图 9- 21：逻辑卷池的配置示例

表 9-2：磁盘卷池配置项的说明

名称	值	含义
存储池名称	Server212-lvm	存储池的英文名称。必须和“目标-路径”中的文件名相同。
存储池类型	逻辑卷池	
逻辑容量	98770	逻辑卷池的总容量大小，应为实际物理磁盘分区容量之和。例如逻辑卷池使用 sda3 和 sda4，则逻辑卷池总容量为 sda3 和 sda4 容量之和。
目标-路径	/dev/server212-lvm	逻辑卷池的路径
源-磁盘卷池-绑定主机	Server212.txy.com	这是选择的，即 LVM 卷组所在的服务器。
存储池格式	lvm2	仅有这种格式。
提供者		这是一个辅助信息，可选择填写
产品名		这是一个辅助信息，可选择填写
设备-路径	/dev/sda3	指定用于构建 LVM 卷组的磁盘分区，可以同时指定多个。

9.3.3 修改存储池描述

可修改选中存储池的描述信息。在系统存储池信息页面，先选中需要修改的存储池，再

点击“描述修改”按钮，弹出描述修改页面，如图 9-7，修改描述信息后点击“确定”按钮保存修改。

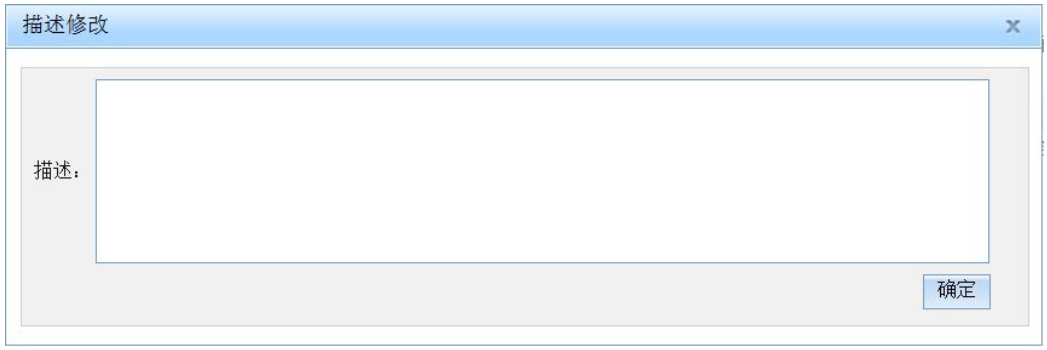


图 9- 7：修改存储池描述信息的界面

9.3.4 删除存储池

在系统存储池信息页面，先选中需要删除的存储池，再点击“删除”按钮，弹出确认删除的页面，最后点击“确定”按钮执行删除操作。*注意，如果此存储池还有存储卷信息，则删除操作将会执行失败。*

9.3.5 查看存储池详细信息

在系统存储池信息页面，先选中需要查看的存储池，再点击“详细信息”按钮，页面跳转至存储池详细信息页面，如图 9-10 所示。

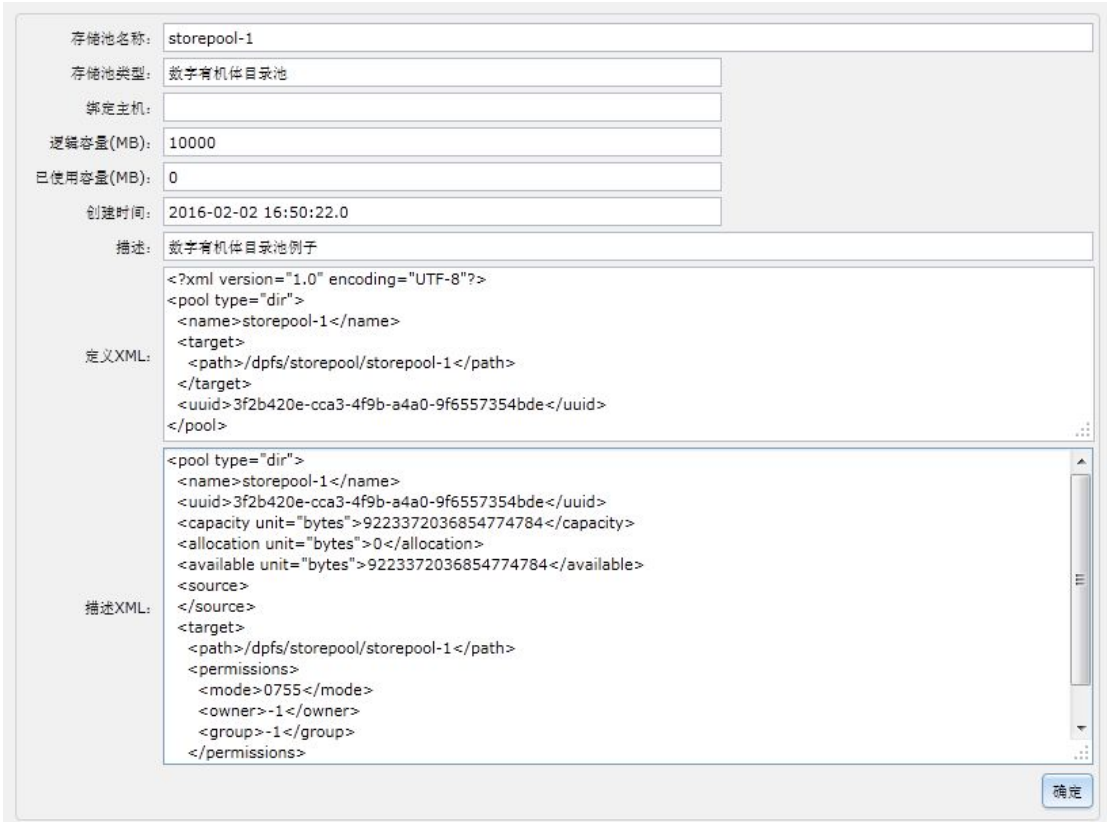


图 9- 8：存储池详细信息显示页面

该页面以 XML 文档的形式描述存储池。该文档的格式以 libvirtd 的规范为准。

9.3.6 存储卷信息

系统管理员也可以对存储池下面的卷进行管理。在系统存储池信息页面，先选中需要管理卷的存储池，再点击“存储卷信息”按钮，页面跳转至存储卷信息页面。由于其具体操作与项目管理员的卷管理是相同的，因此这里不再赘述，详情查看 9.4 章节。

9.4项目管理员管理卷

系统管理员为项目分配可以使用的存储池。项目管理员则负责管理存储池内的卷，并将卷使用在虚拟机上。项目管理员的存储卷管理的内部网页结构如图 9-9。

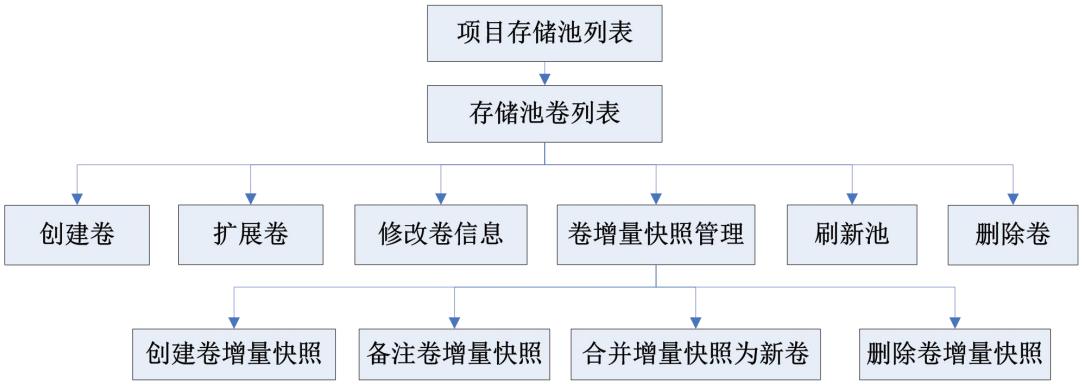
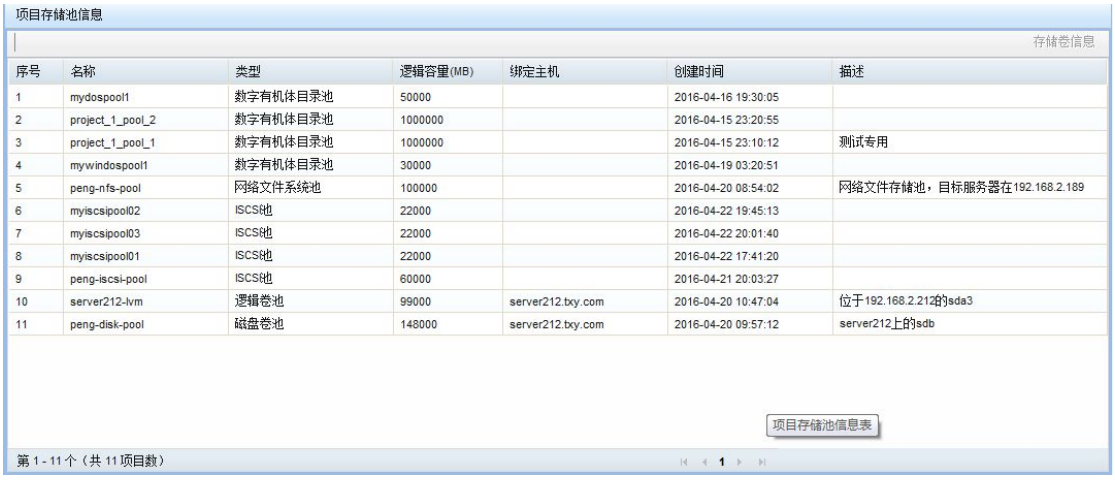


图 9- 9：存储卷管理的页面组织

9.4.1 项目存储池信息

项目管理员的“存储卷管理”的首页面是“项目存储池信息”，如图 9-10。这个页面以列表的形式列出系统的所有存储池。列表顶端只有一个功能按钮：存储卷信息（管理存储池下的卷）。


项目存储池信息

存储卷信息

序号	名称	类型	逻辑容量(MB)	绑定主机	创建时间	描述
1	mydospool1	数字有机体目录池	50000		2016-04-16 19:30:05	
2	project_1_pool_2	数字有机体目录池	1000000		2016-04-15 23:20:55	
3	project_1_pool_1	数字有机体目录池	1000000		2016-04-15 23:10:12	测试专用
4	mywindospool1	数字有机体目录池	30000		2016-04-19 03:20:51	
5	peng-nfs-pool	网络文件系统池	100000		2016-04-20 08:54:02	网络文件存储池，目标服务器在192.168.2.189
6	myiscsipool02	ISCSI池	22000		2016-04-22 19:45:13	
7	myiscsipool03	ISCSI池	22000		2016-04-22 20:01:40	
8	myiscsipool01	ISCSI池	22000		2016-04-22 17:41:20	
9	peng-iscsi-pool	ISCSI池	60000		2016-04-21 20:03:27	
10	server212-lvm	逻辑卷池	99000	server212.txy.com	2016-04-20 10:47:04	位于192.168.2.212的sda3
11	peng-disk-pool	磁盘卷池	148000	server212.txy.com	2016-04-20 09:57:12	server212上的sdb

项目存储池信息表

第 1 - 11 个（共 11 项目数）

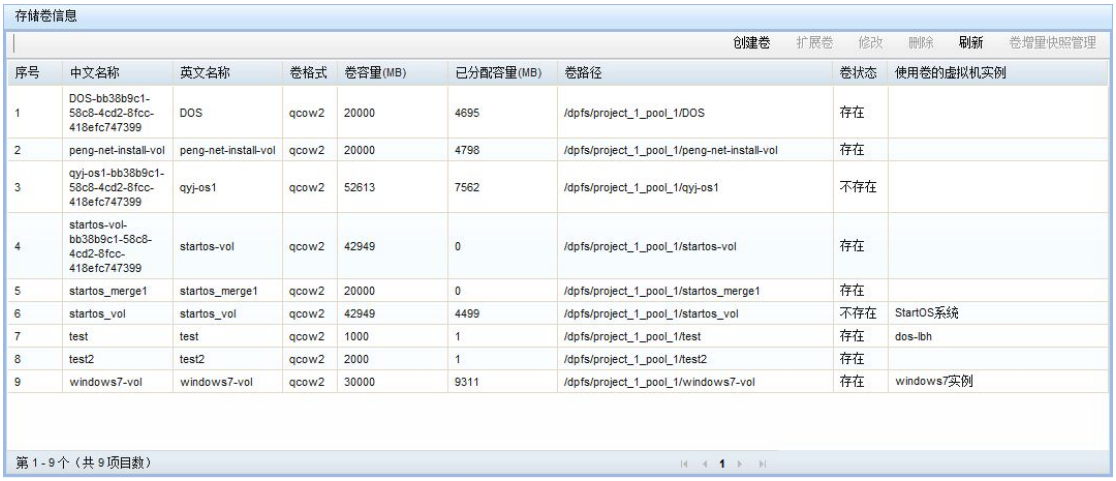
图 9- 10：项目存储池列表信息页面

9.4.2 存储卷信息

可查看存储池下所有卷的信息并对其进行管理。在项目存储池页面上，选择需要管理卷的存储池，再点击“存储卷信息”按钮，页面跳转至存储卷信息页面，如图 9-13。在存储

卷信息页面上，管理员可以进行如下操作：新增卷、扩展卷、修改卷信息、删除卷、刷新池和卷增量快照管理。

若管理员手动的在数字有机体系统中删除文件，或者其他管理人员配置了存储设备，都可能使存储池的卷因非系统的原因而被删除。这是将标准卷状态为“不存在”。



序号	中文名称	英文名称	卷格式	卷容量(MB)	已分配容量(MB)	卷路径	卷状态	使用卷的虚拟机实例
1	DOS-bb38b9c1-58c8-4cd2-8fcc-418efc747399	DOS	qcow2	20000	4695	/dpfs/project_1_pool_1/DOS	存在	
2	peng-net-install-vol	peng-net-install-vol	qcow2	20000	4798	/dpfs/project_1_pool_1/peng-net-install-vol	存在	
3	qyj-os1-bb38b9c1-58c8-4cd2-8fcc-418efc747399	qyj-os1	qcow2	52613	7562	/dpfs/project_1_pool_1/qyj-os1	不存在	
4	startos-vol-bb38b9c1-58c8-4cd2-8fcc-418efc747399	startos-vol	qcow2	42949	0	/dpfs/project_1_pool_1/startos-vol	存在	
5	startos_merge1	startos_merge1	qcow2	20000	0	/dpfs/project_1_pool_1/startos_merge1	存在	
6	startos_vol	startos_vol	qcow2	42949	4499	/dpfs/project_1_pool_1/startos_vol	不存在	StartOS系统
7	test	test	qcow2	1000	1	/dpfs/project_1_pool_1/test	存在	dos-lbh
8	test2	test2	qcow2	2000	1	/dpfs/project_1_pool_1/test2	存在	
9	windows7-vol	windows7-vol	qcow2	30000	9311	/dpfs/project_1_pool_1/windows7-vol	存在	windows7实例

图 9- 11：存储卷管理页面

9.4.3 创建卷

系统只支持在数字有机体目录池、网络文件系统池和宿主机文件目录池下创建卷。这些存储池的卷都是以文件的方式存在，并且由 qemu 模拟器支持。

iSCSI 池、磁盘卷池和逻辑卷池的卷需要在相应的存储系统中配置。iSCSI 池需要在 IP-SAN 存储系统中分配 LUN。磁盘卷池需要用分区工具对磁盘进行分区。逻辑卷池则需要使用 LVM 管理工具在卷组中创建卷。这三类存储池只支持通过“刷新”的方式获得卷信息。刷新存储池时，系统将重新获取实际的卷信息。

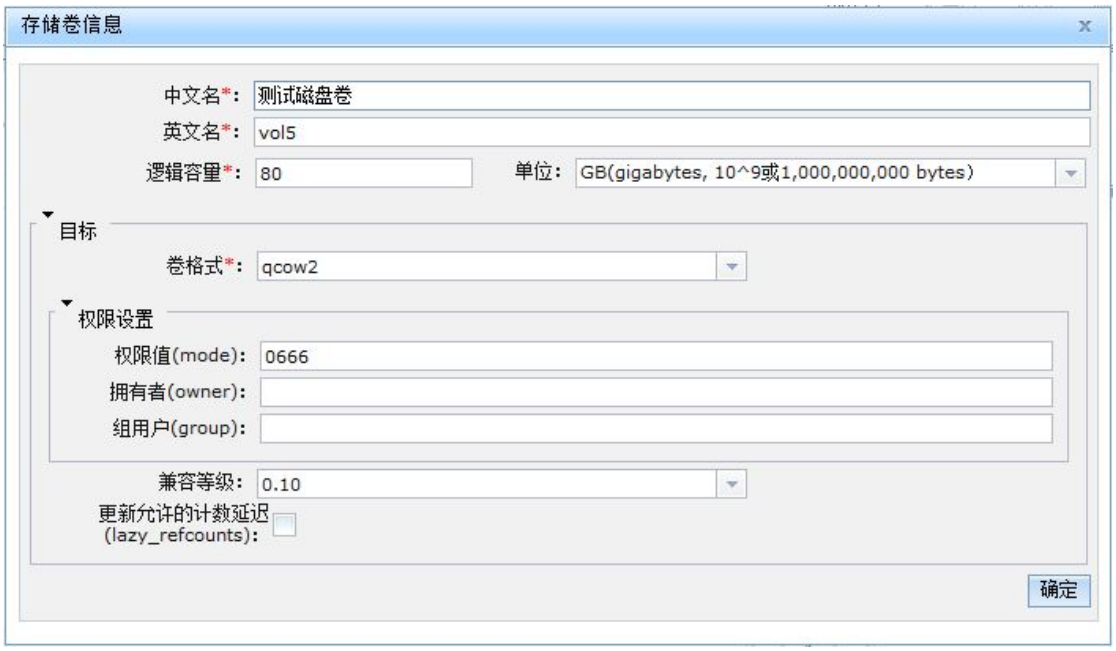


图 9- 12：创建卷的示例

如果存储池支持创建卷，则在存储卷信息页面上，点击“创建卷”按钮，将弹出卷信息的录入面板，如图 9-12。此面板中表单的配置项意义如下：

1) 中文名：方便管理员识别和理解的存储卷名。
2) 英文名：用于创建卷的名称。
3) 逻辑容量：分配给卷的逻辑容量。该值将限制卷逻辑上的最大空间。
4) 目标-卷格式：对于数字有机体目录池、宿主机文件目录池、网络文件系统池来说，这里指的是文件格式，有效值有：qcow2、raw、qcow、qed、cow、vmdk、vpc。
5) 目标-权限设置：提供创建卷时使用的默认权限信息。它包含 4 个子配置。mode 指定 8 进制权限集。owner 指定用户的 ID。group 指定组 ID。lable 指定强制访问控制的 (例如 SELinux) 标签字符串。
6) 目标-兼容等级：指定兼容等级。它只对'qcow2' 类型的卷有效。有效值是 0.10 和 1.1，用于指定映像文件兼容的 QEMU 版本。如果有“feature”元素，则使用 1.1。如果没有指定此值，则使用 qemu-img 的默认版本。
7) 目标-更新允许的计数延迟：更新允许计数延迟。

9.4.4 扩展卷

系统只支持对数字有机体目录池、网络文件系统池和宿主机文件目录池下的卷进行扩容。在存储卷信息页面，先选中需要操作的存储卷，再点击“扩展卷”按钮，弹出扩展卷容量面板，如图 9-13。在面板的文本框内输入需要扩展的卷大小，最后点击“确定”按钮进行保存。

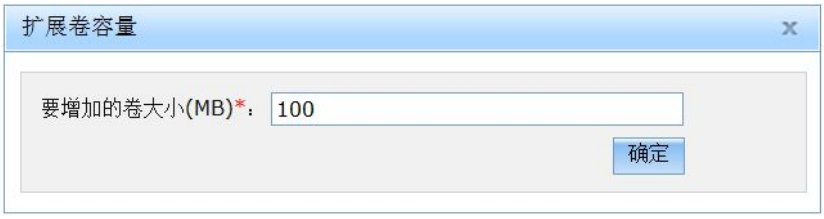


图 9- 13：扩展卷容量的输入界面

扩展的大小必须大于 10MB，且不能缩小存储卷。而且，扩展卷也可能操作失败，其原因可能是空间不足，或者存储卷无法访问。例如存储卷正在被某个虚拟机使用。

9.4.5 修改存储卷

可修改选中存储卷的中文名和描述信息。在存储卷信息页面，先选中需要操作的存储卷，再点击“修改”按钮，弹出卷信息修改面板，如图 9-14 所示。

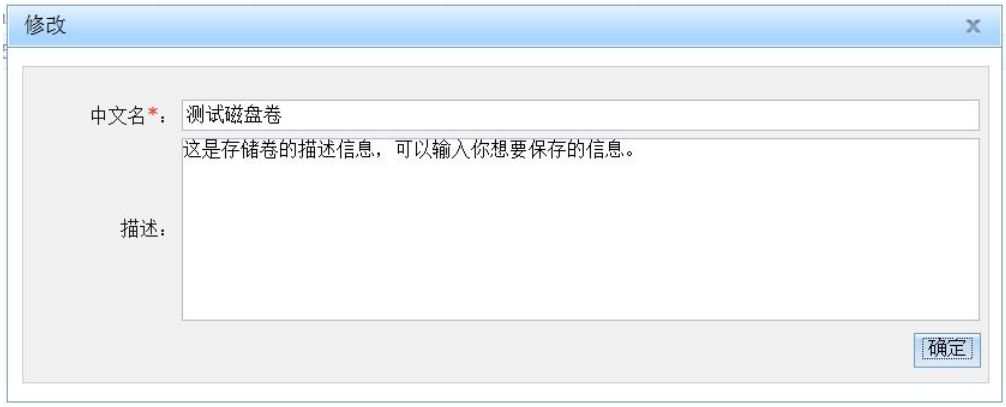


图 9- 14：修改存储卷的中文名和描述信息

9.4.6 删除存储卷

无论存储池是否支持，都可以删除选中的存储卷。对数字有机体目录池、网络文件系统池和宿主机文件目录池来说，删除卷将删除对应的存储文件，以及在数据库中的记录。对 iSCSI 池、磁盘卷池和逻辑卷池来说，删除卷则仅仅删除数据库中的记录，存储卷对应的存储对象需要在想要的存储系统中删除。

如果存储卷正在被虚拟机实例使用，则存储卷是不能被删除的。对其进行删除操作将报错。可以在存储卷浏览页面上看到那个虚拟机实例在使用该存储卷。

在存储卷信息页面，先选中需要删除的存储卷，再点击“删除”按钮，弹出确认删除的页面，最后点击“确定”按钮执行删除操作。注意，如果此存储卷正在被使用，那么删除操作将会执行失败。

9.4.7 刷新存储池

对 iSCSI 存储池、磁盘卷池和逻辑卷池来说，刷新存储池是获得存储卷的唯一方式。而对数字有机体目录池、网络文件系统池和宿主机目录池来说，由于文件可能因外部的操作而被删除，或者手动的在目录下建立的磁盘镜像文件，因此刷新存储池也是获取存储池中正确的卷信息的有效途径。

在存储卷信息页面，点击“刷新”按钮，弹出是否确定刷新的面板，如图 9-18，再点击面板上的“确定”按钮确认操作。

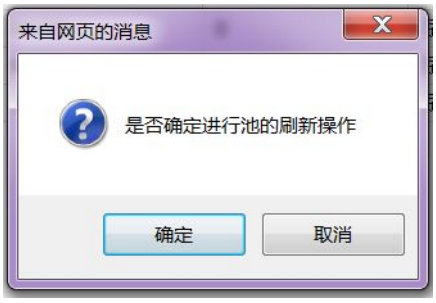


图 9- 15：刷新存储池的确认提示

注意，如果卷由外部程序删除了，则在刷新存储池时将使得对应记录的卷的状态为“不存在”。这时，使用卷的虚拟机可能因卷不存在而故障或者无法启动。项目管理员可以重新指定虚拟机使用的卷，并将状态为“不存在”的卷记录删除。

新获得的卷的中文名称由文件名和存储池的 UUID 构成。可以通过“修改存储卷”来修改中文名称。

9.4.8 卷增量快照管理

9.4.8.1 卷增量快照信息

图 9- 16：存储卷快照管理的界面

对以镜像文件形式存储的卷，可以通过 qemu-img convert 从卷或者卷的某个快照生成出一个新的镜像文件，此镜像文件是一个增量文件。可以根据需要把增量文件的卷文件合并成一个新卷。

卷增量快照的主页面是存储卷快照信息页面。如图 9-16 所示。

9.4.8.2 创建卷增量快照信息

在卷快照信息页面，点击“创建卷增量快照”按钮，弹出存储卷快照信息面板面板，如图 9-17 所示。在面板下的表单中录入新建快照的中文名称和英文名称，最后点击“确定”按钮保存结果。

图 9- 17：创建卷增量快照的输入界面

要注意的是：卷的增量快照也作为一个卷看待。这是为了在创建虚拟机实例时方便选择存储卷的快照作为磁盘设备。但是，如果继续直接使用原有的卷，则快照将失去意义。因此，在为卷创建快照后，应当转而使用卷的快照作为虚拟机的磁盘，并保证卷文件不再被修改。可以为一个卷同时创建多个增量快照。不同的快照给不同的虚拟机使用。这样可以节约存储空间，并共享已经安装好的系统。

9.4.8.3 备注增量快照信息

在卷快照信息页面，先选中需要操作的快照信息，再点击“备注卷增量快照”按钮，弹出存储卷快照信息面板面板，如图 9-18 所示。在面板下的表单中录入备注信息或者中文名称，最后点击“确定”按钮保存结果。

图 9- 18：备注增量快照信息

9.4.8.4 合并增量快照

可合并选中快照文件和卷文件并生成一个新的卷文件。在卷快照信息页面，先选中需要操作的快照信息， 再点击“合并增量快照为新卷”按钮，弹出存储卷信息面板面板，如图 9-19 所示。在面板下的表单中录入新卷的英文名称，最后点击“确定”按钮保存结果。这个操作可能需要较长的时间。需要的时间长短和卷的大小有关。卷越大需要的时间越长。

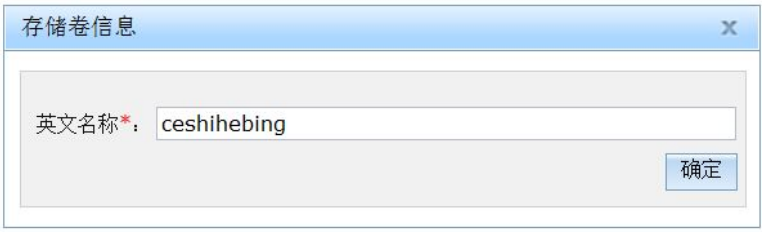


图 9- 19：合并增量快照为新的卷

9.4.8.5 删除卷增量快照

在卷快照信息页面，先选中需要操作的快照信息，再点击“删除卷增量快照”按钮，弹出确认删除的页面，如图 9-20 所示。最后点击“确定”按钮执行删除操作。

图 9- 20：删除卷增量快照的确认对话框

9.5镜像管理

镜像文件指用于模拟磁盘或者光盘设备的数据文件。在数字有机体虚拟机系统中，镜像文件主要指两类文件：

1）用于模拟只读光盘的文件。典型的是操作系统的安装光盘。这类文件采用的是原始（raw）格式。

2）用于运行虚拟机的系统镜像。对许多虚拟机来说，可以将系统盘和数据盘分开，甚至只有系统盘没有数据盘（业务数据在数据库或者网络服务器中）。系统盘除了已经安装好的操作系统和软件外，并不保存其他的永久性数据，因此可以采用临时快照的方式运行。系统盘作为一个只读镜像，给许多虚拟机共享。

大多数镜像文件都是只读的，并且给多个虚拟机共享使用的。如果仅仅是给单个虚拟机使用的文件，建议不作为镜像文件看待。

通常，镜像文件用单个文件模拟一个磁盘或者光驱设备。且必须是 qemu 系统支持的格式。例如光盘的 ISO 镜像，格式为 iso9660（在配置镜像是选择 raw 格式）；用 qemu-img 创建的镜像，格式为 qcow、qcow2、raw 等。其他格式的镜像文件暂不支持。

镜像文件由项目管理员上传，保存到项目组的文件目录中。项目组中的项目都可使用它。如果浏览器不支持文件上传功能，则需要手工通过 ftp 上传到服务器的相应目录，即项目组的文件目录。每台服务器都可以开启 vsftpd 服务，其服务端口为 1144。可以通过各种 ftp 工具上传文件到服务器上，然后拷贝到项目组文件目录下。如果项目组的文件目录在数字有机体文件系统中，则在任意一台服务器上上传一次即可。因此，建议将项目组的文件目录设置在数字有机体文件系统中。

镜像管理的功能主要是查看、新增、删除和修改信息。下面将分别介绍它们。

9.5.1 镜像管理版块的界面

镜像管理页面的组织结构如图 9-21 所示。

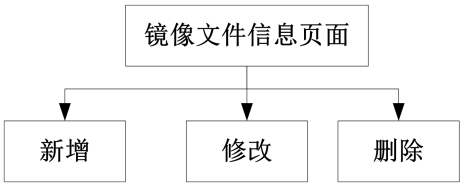


图 9- 21：镜像管理的页面组织

项目管理员登录系统后，点击导航栏中的“镜像管理”按钮就进入镜像文件信息页面。该页面如图 9-22 所示。在镜像文件信息列表的顶端有“新增”、“修改”和“删除”按钮。

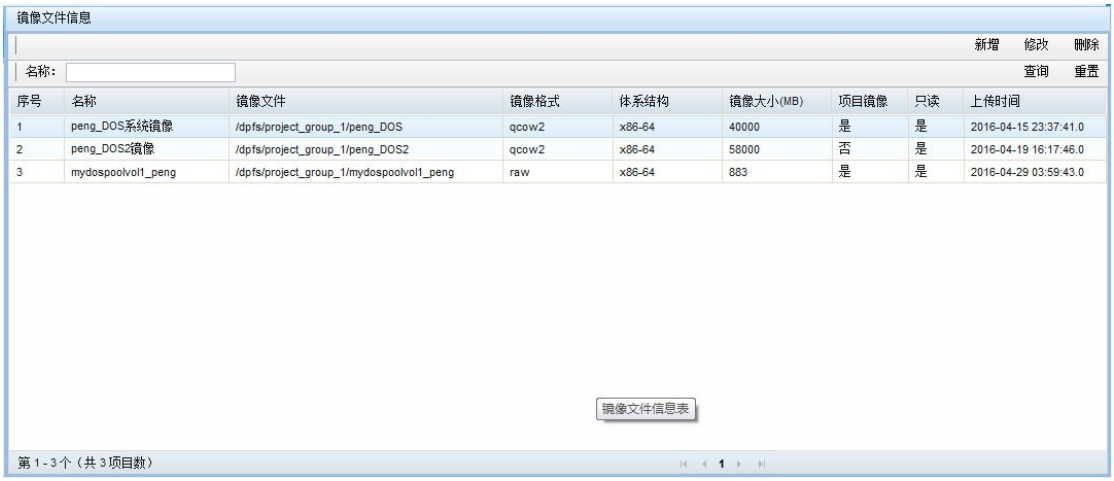


图 9- 22：镜像文件信息页面

9.5.2 新增镜像文件

镜像文件可以直接通过浏览器上传，也可以使用其它方式上传到相应目录中。在数字有机体服务器上，可以用“service vsftpd start”命令启动文件传输服务，服务的端口为 1144。FTP 服务使用操作系统的用户。系统预置有“dos”用户，初始的口令为“123456”。如果修改了口令则使用修改后的口令。在管理终端（即要上传文件的客户机）上可用 ftp 工具将文件上传到服务器。你需要登录到服务器，并将文件移动到项目组文件保存目录下。数字有机体系统有默认用户“dos”，初始口令为“123456”，且 ssh 服务默认是开启的。可以使用 dos 用户登录数字有机体服务器。注意：不能使用“root”用户登录服务器，它被禁止远程登录了。

镜像文件是被项目管理员上传的，项目管理员能查看并使用共处同一项目组的所有镜像文件。并且允许项目管理员修改这些共处同一项目组的所有镜像文件的属性。但是，如果想要删除某个镜像文件，则需要是上传该镜像的所属项目管理员才行，这样有利于防止镜像文件被恶意破坏，能被登录的项目管理员删除的镜像文件的“项目镜像”属性被标记为“是”。

在镜像文件信息页面上，点击“新增”按钮就弹出如图 9-23 所示的新增镜像文件对话框。在对话框中填写镜像文件的信息。各个配置项的填写说明如下：

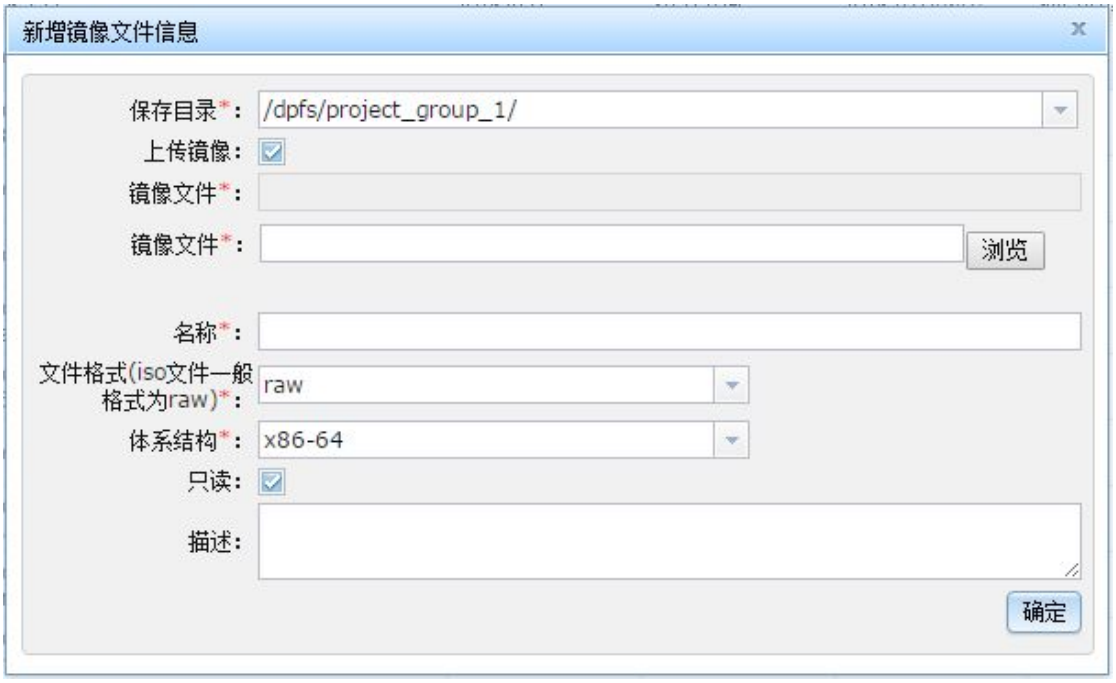


图 9- 23：新增镜像文件的对话框

1) 名称：镜像文件的中文名称，方便记录和辨认。
2) 文件格式：镜像文件的处理格式。以 iso 为后缀的光盘镜像文件需要选择 raw 格式。其他格式的镜像文件可以用“qemu-img info 文件名”命令来查看文件的格式，即显示的“file format”。如果格式选择错误，镜像文件将无法正确使用。
3) 体系结构：镜像文件适合的虚拟机类型。该项目还未用作配置限制，仅用于配置虚拟机时提示用户。
4) 虚拟机间是否共享：指镜像文件是否要在虚拟机间共享。
5) 是否只读：指镜像文件是否是只读的。如果镜像文件是只读的，在被用于模拟可写磁盘时，系统将自动为其创建一个临时快照，确保镜像文件本身不被修改。非只读的镜像文件则无此处理。
6) 描述：镜像文件的描述信息，根据需要输入。
7) 保存目录：每个项目都属于某些项目组（至少属于一个项目组）。项目上传的镜像文件保存在所属项目组的文件保存目录下。当项目属于多个项目组时，可以选择保存的目录。
8) 是否上传文件：指是否要通过浏览器上传文件。如果不上传文件，则必须输入文件名称。如果浏览器不支持文件上传功能，则用户需要使用其它方式上传，如 ftp。
9) 文件名称：已经上传的文件的名称，它相对于保存目录。
10) 镜像文件：当需要上传文件时，可通过其后的浏览按钮选择本地要上传的文件。

如果不需要上传文件，则点击“确定”时后台直接保存镜像文件记录。如果要上传文件，则需要等待文件上传完成才能保存镜像文件记录，因此需要较长时间的等待。

9.5.3 修改镜像文件信息

在镜像文件信息页面上选择好要修改的镜像文件记录后，点击工具栏中的“修改”按钮就弹出修改镜像文件信息的对话框，其界面如所示。

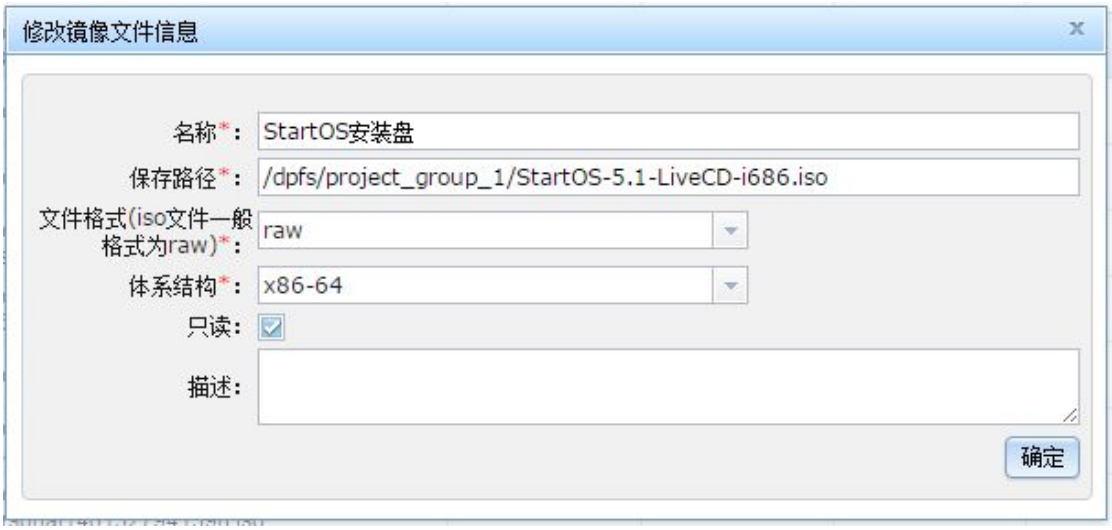


图 9- 24：修改镜像文件信息的对话框

除了保存路径外，该对话框上的所有配置项都可以修改。各个配置项的含义和新增镜像文件对话框中的含义是相同的，这里不再描述。修改完成后，点击“确定”按钮就保存修改后的信息。

9.5.4 删除镜像文件

在镜像文件信息页面上选择好要删除的镜像文件记录后，点击工具栏中的“删除”按钮就弹出确认是否删除的对话框。如果点击“确定”则镜像文件将被删除。如果镜像文件还被某个虚拟机实例使用，则删除操作将失败。而且，即使使用镜像文件的虚拟机实例没有运行也一样要失败。

要注意的是：删除镜像文件不仅会删除数据库中的记录，也会删除真实的文件。请谨慎操作。

10.网络虚拟化管理

10.1 概述

有关网络虚拟化的方案及部署请参见第三章的描述。总体上讲，虚拟网络的部署是和物理网络相关的，并和物理网络紧密结合在一起。当虚拟网络需要和外界通信时，就涉及到和物理网络的互联问题。

在数字有机体虚拟机系统中，虚拟网络被划分为四类，即隔离网络、NAT 网络、路由网络和桥接网络。前三类网络其实可以归为一类，即单宿主机内的内部虚拟网络。它们都通过宿主机内的一个 Linux 桥接器来模拟网络 HUB，且这个桥接器并不直接关联宿主机的物理网络接口。因此，虚拟机的消息并不能直接转发到物理网络上。虚拟机的虚拟网络接口用一个 tap 设备（Linux 下的一种虚设备）模拟，tap 设备直接连接到桥接器上，因此同一个桥接器上的虚拟机可以相互通信。当虚拟机需要和外部网络通信时，必须由宿主机的协议栈进行转发。宿主机只转发虚拟网络的 IP 包，而不是转发链路上的帧，因此虚拟网络和外界网络是通过 IP 层互联的。消息转发的方式有 NAT（网络地址转换）和路由，因此又分为 NAT 网络和路由网络，而不与外界网络通信的内部虚拟网络就是隔离网络了。

桥接网络根据实现的方式不同，可以分为三类，即 Linux 桥接、ovs 桥接和 macvtap 桥接网络。和单宿主机内的内部虚拟网络不同，这三类桥接都将宿主机的物理网络接口和桥接器（或者内核的帧队列）直接关联起来，在链路层转发桥接器端口上的消息。换句话说，宿主机的物理网络接口也被作为桥接器的一个端口。因此，外部网络通过宿主机的物理接口直接连接了虚拟网络的 HUB，这样虚拟网络也就在链路层接入了外部网络。如果外部网络是一个交换网，则虚拟网络就是这个交换网的一部分。因此，桥接网络的虚拟网络和外部网络是融合在一起的。桥接网络可以跨宿主机组建，即不同宿主机间的虚拟机可以通过物理交换网络直接在链路层通信。

理解了这些虚拟网络的实质后，即可根据需要选择要部署的虚拟网络的类型。具体内容参见第三章的描述。

不过，就虚拟网络管理来说，还需要从网络规划的角度来考虑虚拟网络的部署。当前，IP 交换网络是最常见的局域网组织形式。规划 IP 交换网络的思路是：

1）分析业务的网络需求，尤其是业务的虚拟机的通信需求，理清哪些网络实体（虚拟机的每个网络接口）间需要通信，以及通信的带宽、服务质量要求。

2）根据业务的通信需求，规划系统的子网划分。通常将需要紧密通信的，或者需要相同网络安全的网络实体划分在同一个子网中。子网和子网间通过路由器互联。

3）规划每个子网的实现方式。如果子网内的虚拟机需要跨宿主机，则选择桥接网络实现，否则根据子网和外部通信的方式选择 NAT 或者路由网络。

4）规划子网的内部服务。系统可以为子网提供 DHCP、tftp、bootp 和 DNS 服务。

5）根据子网的部署，规划子网间的互联方案，即规划网络路由。重点是网关的配置。

6）根据每个通信实体（这里指虚拟机网络接口）网络安全需求，涉及安全组。

7）根据规划配置虚拟网络。

在数字有机体虚拟机系统中，配置虚拟网络的思路和规划 IP 交换网络的思路是相同的。系统管理员根据子网规划，配置系统要使用的子网，包括配置每个子网的内部服务、子网的网关等。要注意的是，隔离网络、NAT 网络和路由网络的网关都由虚拟网络所在的宿主机担任，因此子网的内部服务也由其提供。桥接网络的网关既可以由某台宿主机担任，也可以由物理路由器担任，因此子网内部服务也可以由某台宿主机或者物理路由器担任。

由于不同的子网可能需要相同的网络安全，加上网络过滤规则难以设计，因此将安全组管理交给系统管理员负责。系统管理员根据通信实体的安全需求设计安全规则（即过滤器），然后将过滤器组合为安全组。项目管理员则根据虚拟机的网络接口安全需求选用安全组。

项目管理员根据虚拟机的网络需求，建立虚拟网络。虚拟机的网络接口都必须关联某个虚拟网络。虚拟网络可以关联子网，也可以不关联子网。如果关联子网，则可以认为网络内的虚拟机接口使用该子网的 IP 地址、内部服务和网关。如果不关联子网，则要求虚拟机使用者自己配置虚拟机的网络接口。

在配置虚拟机的网络接口时，项目管理员选择网络接口连接的虚拟网络，同时也可以给网络接口指定安全组。

10.2 系统管理员管理子网

10.2.1 子网管理

这里的子网指 IP 子网。互联网用 IP 地址来标识通信者的地址。IP 地址是一个正整数。其长度有 32 位（IPV4）和 128 位（IPV6）两种。以 8 位为一个字节，则有 4 字节地址和 16 字节地址两种。每个字节可以用 0 到 255 的整数记录，则 IPV4 的地址可以表示为 4 个 0 到 255 的数字，数字间用点分隔，这就是 IPv4 地址的点分格式，例如 192.168.2.7。IPv6 的地址则不用点分格式表示。

互联网以 IP 路由为基础。按照 TCP/IP 协议的结构，物理链路层可以相互通信的实体组织为一个子网，用某段 IP 地址来表示他。为了方便，将 IP 地址拆分为两段，一段用于区分不同的子网，称作网络地址，余下的位用于区分子网内的通信实体，称作主机号。网络地址根据预留给主机号的位数不同，可以分为 A 类（24 位）、B 类（16 位）和 C 类（8 位）三种。D 类地址是 C 类中再划分出的一部分，用于表示组播地址。显然一个 C 类子网可以有 255 台主机（0 号地址被保留作为网络地址）。但是，有些时候子网的主机数比 255 小得多，为了充分利用地址，又允许在表示主机号的位中再划出部分位，用于表示子网的号，将表示子网的位和表示网络地址的位合并，就形成了子网的网络地址。如果将用于表示网络地址的位设为 1，其余位设为 0，就得到了一个掩码。将掩码和通信主机的 IP 地址相与，就得到了主机 IP 地址所在的子网的网络地址。通常将地址的高位留给网络地址，因此可以直接用留给网络地址的位数来表示掩码。例如，某个 IP 子网表示为 192.168.2.192/26，则表示其网络地址为 192.168.2.192，网络掩码是 26，子网的 IP 地址范围为 192.168.2.192 到 192.168.2.255。当然，也可以用不连续的位表示子网，不过这要更复杂些。

为了充分利用 IP 地址，或者为了方便分配 IP 地址，系统支持 DHCP 协议，即动态主机地址分配协议。同时，一个子网内的主机通常有某些共同的服务需求，例如域名解析（DNS）、网络启动服务（bootp）和 tftp 服务等。可以将这些服务和子网关联起来，将子网看做一种资源，将其分配给项目。项目再利用子网来建立虚拟网络。这就是虚拟机系统中 IP 子网的另一层含义。

数字有机体虚拟机系统的子网管理的页面内部结构如图 10-1 所示。

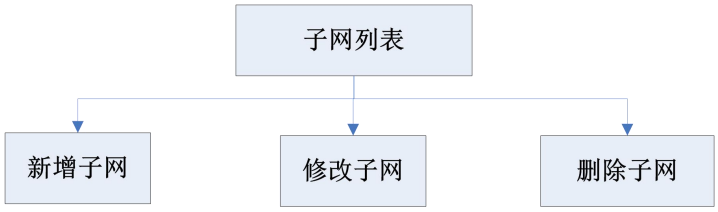
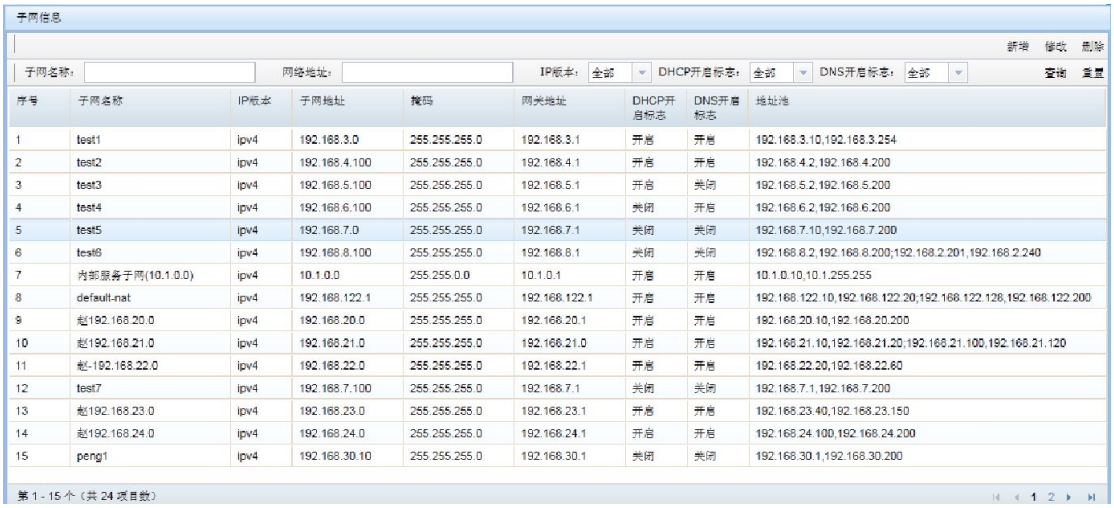


图 10- 1：子网管理的内部结构图

10.2.1.1 子网信息

系统管理员的网络管理的首页是子网信息页面，如图 10-2。这个页面以列表的形式列出系统的所有子网。系统管理员可以根据子网名称、网络地址、IP 版本、DHCP 开启标志、DNS 开启标志等条件过滤数据。数据结果分页显示。列表顶端只有三个常用功能按钮：新增、修改、删除。


子网信息

新增 修改 删除

子网名称： 网络地址： IP版本： 全部 DHCP开启标志： 全部 DNS开启标志： 全部 查询 重置

序号	子网名称	IP版本	子网地址	掩码	网关地址	DHCP开启标志	DNS开启标志	地址池
1	test1	ipv4	192.168.3.0	255.255.255.0	192.168.3.1	开启	开启	192.168.3.10,192.168.3.254
2	test2	ipv4	192.168.4.100	255.255.255.0	192.168.4.1	开启	开启	192.168.4.2,192.168.4.200
3	test3	ipv4	192.168.5.100	255.255.255.0	192.168.5.1	开启	关闭	192.168.5.2,192.168.5.200
4	test4	ipv4	192.168.6.100	255.255.255.0	192.168.6.1	关闭	开启	192.168.6.2,192.168.6.200
5	test5	ipv4	192.168.7.0	255.255.255.0	192.168.7.1	关闭	关闭	192.168.7.10,192.168.7.200
6	test6	ipv4	192.168.8.100	255.255.255.0	192.168.8.1	关闭	关闭	192.168.8.2,192.168.8.200;192.168.2.201,192.168.2.240
7	内部服务子网(10.1.0.0)	ipv4	10.1.0.0	255.255.0.0	10.1.0.1	开启	开启	10.1.0.10,10.1.255.255
8	default-nat	ipv4	192.168.122.1	255.255.255.0	192.168.122.1	开启	开启	192.168.122.10,192.168.122.20;192.168.122.126,192.168.122.200
9	赵192.168.20.0	ipv4	192.168.20.0	255.255.255.0	192.168.20.1	开启	开启	192.168.20.10,192.168.20.200
10	赵192.168.21.0	ipv4	192.168.21.0	255.255.255.0	192.168.21.0	开启	开启	192.168.21.10,192.168.21.20;192.168.21.100,192.168.21.120
11	赵-192.168.22.0	ipv4	192.168.22.0	255.255.255.0	192.168.22.1	开启	开启	192.168.22.20,192.168.22.60
12	test7	ipv4	192.168.7.100	255.255.255.0	192.168.7.1	关闭	关闭	192.168.7.1,192.168.7.200
13	赵192.168.23.0	ipv4	192.168.23.0	255.255.255.0	192.168.23.1	开启	开启	192.168.23.40,192.168.23.150
14	赵192.168.24.0	ipv4	192.168.24.0	255.255.255.0	192.168.24.1	开启	开启	192.168.24.100,192.168.24.200
15	peng1	ipv4	192.168.30.10	255.255.255.0	192.168.30.1	关闭	关闭	192.168.30.1,192.168.30.200

第 1 - 15 个（共 24 项目数） 1 2

图 10- 2：子网信息列表

10.2.1.2 新增子网

在子网信息页面，点击“新增”按钮，页面跳转至子网配置页面，如图 10-3。此页面中有三个子面板：子网基本信息、DHCP 服务配置、DNS 服务配置。默认只开启了子网基本信息子面板。只有选中子网基本信息子面板中的启用 DHCP 服务和启用 DNS 服务，才能开启对应的子面板。下面分别介绍三个子面板：

子网基本信息配置如图 10-3 所示。各项配置的含义如下：

1）子网名称：便于理解和记忆的子网名称。名称可由字母、数字、中文、中划线、下划线和点组成。

图 10- 3：新增子网之子网基本信息

2）子网描述：对当前子网的详细说明，仅用于方便用户对此子网的了解。

3）IP 地址：子网相关的 IP 地址信息。这是一组信息的集合。它包括以下几项内容：

（1）IP 地址-IP 版本：指定地址类型，有效值有：ipv4 、ipv6。现在主要支持 IPv4。

（2）IP 地址-子网地址：子网的网络地址，即子网内 IP 地址的共同前缀。将子网

内的 IP 地址和网络掩码相与就得到网络地址。

（3）IP 地址-网络掩码：对 IPv4 地址，此属性以点-数字的格式定义网络地址中的有效位。对 IPv6 地址，此属性使用前缀位数的方式指定，它是一个整数。

（4）IP 地址-网关地址：网关是子网和外部网络通信的路由器或者完成路由的主机的地址。网关地址必须是子网内的一个地址，这样子网内的其他地址可以通过链路层到达网关。如果没有其他路由信息，子网内的主机都将无法直接送达的 IP 包发送给网关，由网关进行转发。

（5）IP 地址-绑定网关到宿主机：如果需要系统自动建立网关，则可以设定网关所在的宿主机。

4）地址池：这里指子网中可以使用的 IP 地址范围。如果采用 DHCP 方式分配地址，则是 DHCP 服务可以分配的地址范围。即使不采用 DHCP 服务，当在子网的网络上部署虚拟机集群时，地址池也用于指定那些地址是合法的 IP 地址。地址池下有新增、删除两个按钮，可以对表格中的地址范围进行增删操作，新增面板如图 10-4。

5）启用 DHCP 服务：是否需要网关为子网提供 DHCP 服务。如果需要则开启 DHCP 服务配置面板，可对 DHCP 服务进行配置。

6）启用 DNS 服务：是否需要网关为子网提供 DNS 服务。如果需要则开启 DNS 服务配置面板，可对 DNS 服务进行配置。

7）TFTP 根路径：TFTP 服务的根路径，TFTP 服务是一个传输文件简单协议，它可用于 UDP 环境且其代码占用的内存较小，但它仅支持 IPv4。

8）需要自动建立网关：由系统在绑定的宿主机或随机宿主机上自动建立子网的网关。同时可以配置启动和关闭网关时的执行脚本。其中，隔绝网络、NAT 网络和路由网络需要将此选项设置为不自动建立。

9）启动网关时执行的 shell 脚本：配置系统自动建立的网关启动时的执行脚本。

10）关闭网关时执行的 shell 脚本：配置系统自动建立的网关关闭时的执行脚本。

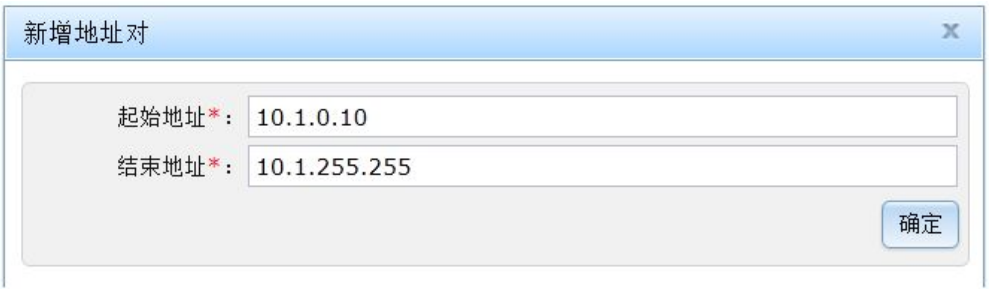


图 10- 4：新增子网之新增地址对

IP 子网的 DHCP 服务配置界面如图 10-5 所示。各项参数的含义是：

1）BOOTP 参数：指定 DHCP 协议服务的 BOOTP 参数，仅支持 IPv4。“启动文件”指虚拟机通过网络启动时访问的境像文件名称，“获取地址”指获取境像文件的 TFTP 服务器地址。当同时配置有“TFTP 根路径”时，默认的 TFTP 服务器就是 DHCP 服务器。

2）IP 预定列表：实现 IP 到主机的绑定。面板上有新增、删除两个按钮，可以对表格中的记录对进行增删操作。新增面板如图 10-6。IPv4 的主机用 MAC 地址标识。MAC 地址是

网络接口的媒体访问地址（链路层的地址）。如果配置了主机名称，那么当前主机名称将覆盖原主机名称。IPv6 的主机使用主机名称标识，主机名称是客户机发送给服务器的自己的普通名字。

图 10- 5：新增子网之 DHCP 服务配置

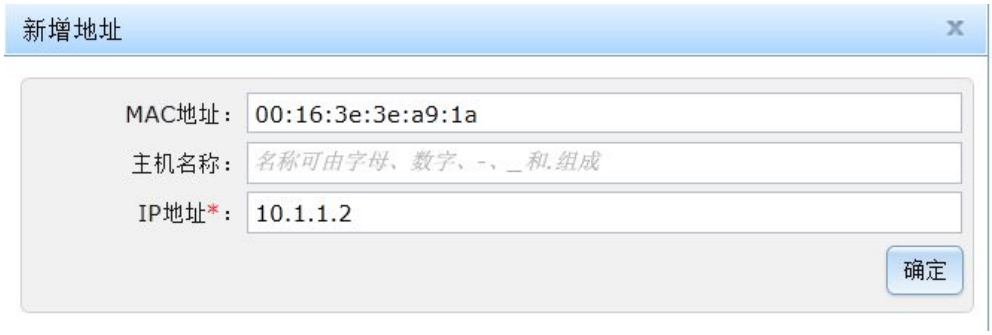


图 10- 6：新增子网之新增 IP 预定地址

子网的域名解析服务（DNS）配置如图 10-7、图 10-9 和图 10-12 所示。配置说明如下：

图 10- 7：新增子网之 DNS 服务配置一

1）网络域名：定义子网内主机的网络域名。主机的域名为主机名加网络域名。

2）进一步查询简短名字：如果未选中此项，则那些没有 DNS 域名（即未包含网络域名）的解析请求将不转发给上级 DNS 服务器。这时，这样的名字仅在虚拟网络自己的 DNS 服务器能解析时才被解析。如果选中此项，则要转发。

3）名字列表：DNS 查询列表。面板下有新增、删除两个按钮，可以对表格中的记录对进行增删操作。新增面板如图 10-8 所示。名字是通过 DNS 可查询的名字，值是该名字被查询时返回的值。名字不能包含空格或者逗号。值是一个单字符串，可包括多个用分号分隔的 IP 地址值。

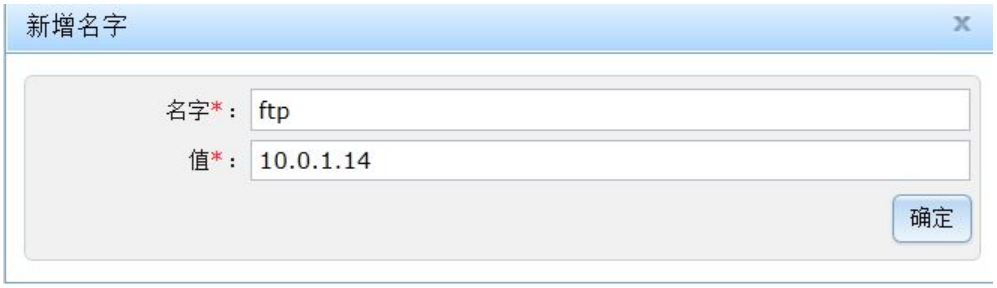


图 10- 8：新增子网之新增名字

4）上级域名服务器列表：上级 DNS 服务器的 IP 地址，如图 10-9。面板上有新增、删除两个按钮，可以对表格中的记录对进行增删操作。新增面板如图 10-10。服务器地址即是 DNS 服务器地址。

图 10- 9：新增子网之 DNS 服务配置二

图 10- 10：新增子网之新增上级服务器

5）服务列表：用于定义提供特定服务的服务器地址，如图 10-9 所示。面板上有新增、删除两个按钮，可以对表格中的记录对进行增删操作。新增面板如图 10-11。面板字段的含义意如下：

1) 服务名：服务名称。
2) 域名：提供服务的主机域名。
3) 目标地址：提供服务的主机 IP。
4) 协议：服务协议名称，有效值为 tcp、udp。
5) 端口： 提供服务的端口，只有目标地址不为空时才有效。
6) 优先级：相同服务的优先选择的参数，只有目标地址不为空时才有效。
7) 权重：相同服务的优先选择的参数，只有目标地址不为空时才有效。

图 10- 11：新增子网之新增服务

6）主机列表：子网内主机名的列表。实现主机域名和 IP 地址的映射，如图 10-12 所示。面板上有新增、删除两个按钮，可以对表格中的记录对进行增删操作。新增面板如图 10-13 所示。

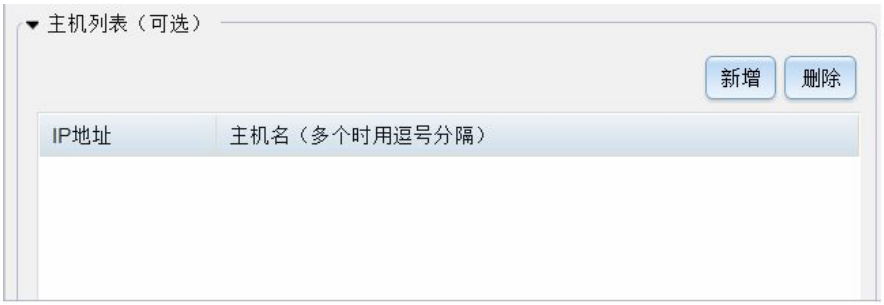


图 10- 12：新增子网之 DNS 服务配置三

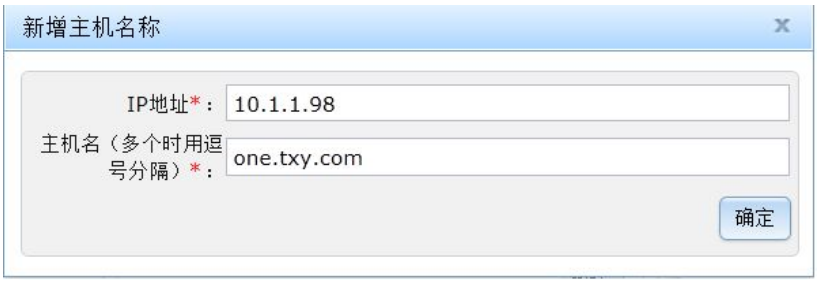


图 10- 13：新增子网之新增主机名称

10.2.1.3 修改子网

在子网信息页面，先选中需要操作的子网，再点击“修改”按钮，页面跳转至子网配置页面。子网是网络的一部分，在新建网络时被引用，修改子网之后，引用旧的子网的网络不会随着子网的修改而自动修改，用户需要重新执行一次“修改网络”才会更新对子网的修改。该界面与操作都与新增子网一致，因此这里就不再赘述。详情参考 10.2.1.2。

10.2.1.4 删除子网

在子网信息页面，先选中需要删除的子网，再点击“删除”按钮，弹出确认删除的页面，如图 10-14，最后点击“确定”按钮执行删除操作。注意，如果此子网正在被使用，那么删除操作将会执行失败。

图 10- 14：删除子网

10.2.2 子网的项目配置

配置好子网后，由系统管理员将子网分配给项目。一个项目可以使用多个子网，一个子网也可以被多个项目共享。要注意一些特殊的子网：比如 192.168.x.x 这样的 C 类子网和 10.x.x.x 这样的 A 类子网。它们被用于表示内部通信子网，即不通过路由和互联网络通信的子网。这类地址被称为私有地址，它们常被用作隔离网络或者 NAT 网络的子网地址。

图 10- 15：子网的项目配置信息列表

10.2.2.1 子网的项目配置信息

点击网络管理左导航栏中的“子网的项目配置”项，进入子网的项目配置信息页面，如图 10-15 所示。这个页面以列表的形式列出系统的所有子网的项目配置信息。系统管理员可以根据子网名称、网络地址、项目名称等条件过滤数据。数据结果分页显示。列表顶端只有两个功能按钮：新增、删除。

10.2.2.2 新增子网的项目配置

在子网的项目配置信息页面，点击“新增”按钮，弹出新增可用该子网的项目面板，如图 10-16 所示。选择对应的子网和项目后，点击“确定”按钮提交操作结果。

图 10- 16：新增子网的项目配置

10.2.2.3 删除子网的项目配置

在子网的项目配置信息页面，先选中需要删除的记录，再点击“删除”按钮，弹出确认删除的面板，如图 10-17 所示。最后点击“确定”按钮执行删除操作。

图 10- 17：删除子网的项目配置

10.2.3 过滤器管理

为了限制或者增强虚拟机网络的安全，可以给虚拟机网络接口设置安全组。安全组由一个过滤器实现（该过滤器又可以调用其他过滤器，因此该过滤器被称作顶层过滤器）。过滤器由一组包过滤规则实现，并在应用到虚拟机网络接口时，由系统转换为宿主机上的一系列 IP-Table 规则。如何定义过滤器需要参加 libvirt 的文件。因为这部分的内容过多。

系统中所有过滤器的英文名称不能重复。管理网站将检查英文名称是否重复，并通知项目管理员存在冲突。过滤器的英文名称将在其他过滤器中引用，因此不能随意修改。但是，既然管理网站没有检查引用的正确性，因此也就只能由项目管理员自己来保证了。

项目安全组的过滤器是全系统共享的，系统管理员和项目管理员都可以定义过滤器。过滤器管理的内部结构如图 10-18 所示。

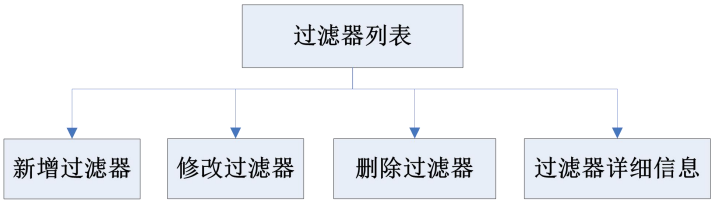


图 10- 18：过滤器管理内部结构图

10.2.3.1 过滤器信息

点击网络管理左导航栏中的“过滤器信息”项，进入过滤器信息页面，如图 10-19 所示。这个页面以列表的形式列出系统的所有过滤器信息。系统管理员可以根据英文名、中文名等条件过滤数据。数据结果分页显示。列表顶端只有四个常用功能按钮：新增、修改、删除、详细信息。

过滤器信息

新增 修改 删除 详细信息

英文名: 中文名: 查询 重置

序号	中文名	英文名	创建日期	创建账号	描述
1	允许ipv4进入	allow-incoming-ipv4	2015-12-08 16:00:48.0	admin	内置过滤器
2	允许DHCP服务	allow-dhcp-server	2015-12-08 15:59:26.0	admin	内置过滤器
3	允许dhcp	allow-dhcp	2015-12-08 15:58:05.0	admin	内置过滤器
4	允许ARP	allow-arp	2015-12-08 15:56:19.0	admin	内置过滤器
5	防IP洪泛	no-ip-spoofing	2015-11-16 15:12:34.0	admin	dsfdfdss
6	过滤器一	no-arp-mac-spoofing	2015-08-29 17:50:41.0	admin	filter1 xml
7	允许ipv4	allow-ipv4	2015-12-08 16:01:48.0	admin	内置过滤器
8	clean-traffic	clean-traffic	2015-12-08 16:02:53.0	admin	内置过滤器

第 1 - 14 个（共 14 项目数） 1

图 10- 19：过滤器信息列表

10.2.3.2 新增过滤器

在过滤器信息页面，点击“新增”按钮，弹出新增过滤器面板，如图 10-20 所示。面板中的“过滤器定义 XML”和 XML 描述可以择一使用。前者是选择已经书写好的 XML 文件（在客户机上，通过浏览本地文件系统选择，系统自动上传），后者则是直接在文字域内录入过滤器定义 XML。

10.2.3.3 修改过滤器

在过滤器信息页面，先选中需要操作的过滤器信息，再点击“修改”按钮，弹出修改过滤器面板，如图 10-21 所示。对于过滤器定义 XML，新增时不管是上传的文件还是直接录入的 XML，修改时都会把内容读取出来直接操作。

新增过滤器

中文名*： 允许DHCP服务

英文名*： allow-dhcp-server

描述： 内置过滤器

过滤器定义XML： allow-dhcp-server.xml 浏览...

XML描述：

保存

图 10- 20：新增过滤器

图 10- 21：修改过滤器

10.2.3.4 删除过滤器

在过滤器信息页面，先选中需要删除的过滤器信息，再点击“删除”按钮，弹出确认删除的面板，如图 10-22 所示。最后点击“确定”按钮执行删除操作。

图 10- 22：删除过滤器

10.2.3.5 过滤器详细信息

在过滤器信息页面，先选中需要查看的过滤器信息，再点击“详细信息”按钮，弹出过滤器详细信息面板，如图 10-23 所示。该页面上只能查看信息，不能修改。

图 10- 23：过滤器详细信息

10.2.4 安全组管理

安全组是一系列过滤器的集合。其中，顶层过滤器引用组中的其他过滤器。其他过滤器又可以相互引用。一个安全组的过滤器集合需要满足以下两个条件：

1）所有集合中的过滤器引用的过滤器都在集合中，即集合的完备性。

2）除顶层过滤器外，不存在不被顶层过滤器递归引用的过滤器，即所有过滤器都是必须的。

要注意的是：管理系统并不检查安全组的过滤器集合是否满足这两个条件。项目管理员需要自己来确保这一点。

除了 libvirt 预定义的过滤器外，安全组中的每个过滤器都需要预先定义。

虚拟机的每个网络接口可以指定一个安全组。可以给虚拟机的不同网络接口指定不同的安全组。安全组的顶层过滤器是网络接口使用的过滤器。

安全组管理内部结构如图 10-24 所示。

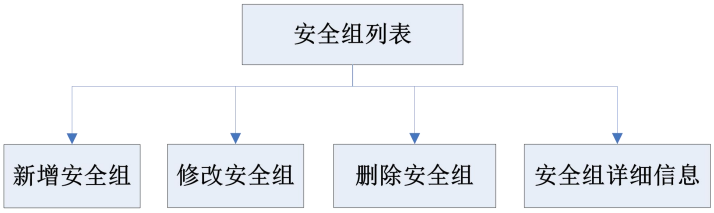


图 10- 24：安全组管理内部结构图

10.2.4.1 安全组信息

点击网络管理左导航栏中的“安全组信息”项，进入安全组信息页面，如图 10-25 所示。这个页面以列表的形式列出系统的所有安全组信息。系统管理员可以根据安全组名称过滤数

据。数据结果分页显示。列表顶端只有四个常用功能按钮：新增、修改、删除、详细信息。

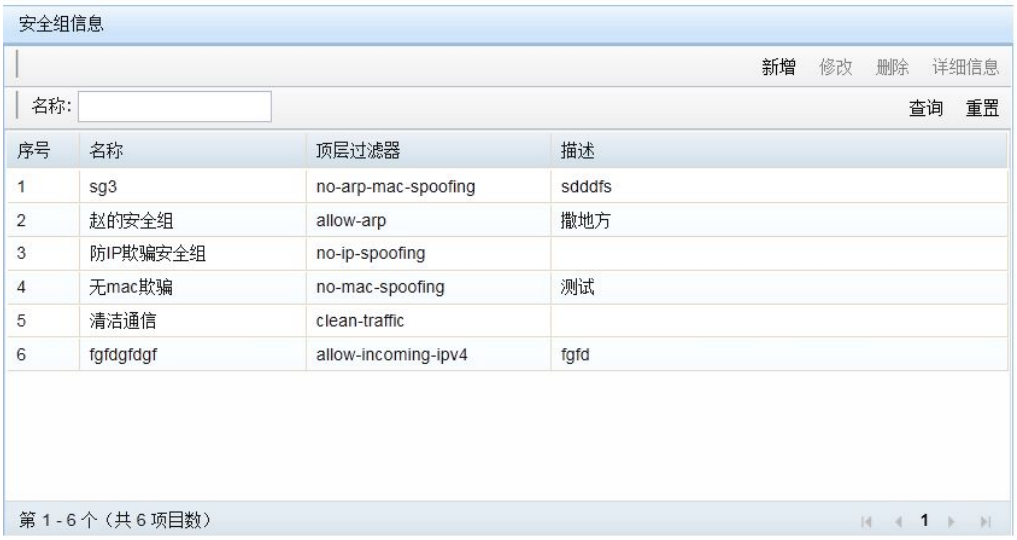


图 10- 25：安全组信息列表

10.2.4.2 新增安全组

在安全组信息页面，点击“新增”按钮，弹出新增安全组面板，如图 10-26 所示。面板中的表单项意义如下：

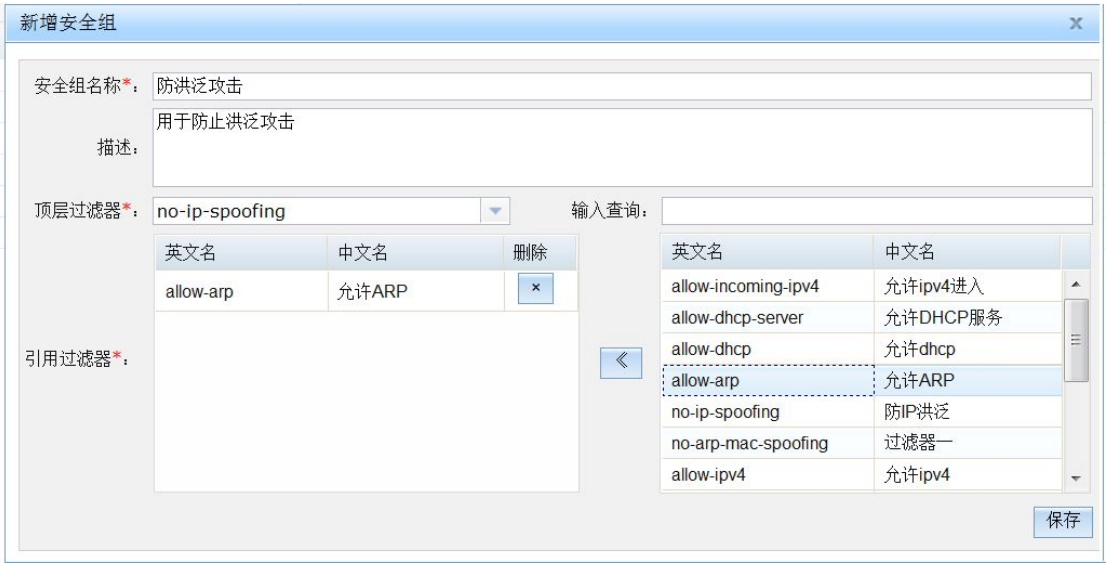


图 10- 26：新增安全组

1）安全组名称：方便管理员识别理解的安全组名。

2）描述：对当前安全组的详细说明，仅用于方便用户对此安全组进一步了解。

3）顶层过滤器：安全组内部的最上层过滤器，即所有引用的过滤器都被此过滤器递归引用。该过滤器的名字将在定义虚拟机网络接口时引用。

4）引用过滤器：顶层过滤器递归引用的所有过滤器。注意，顶层过滤器和引用过滤器不被系统约束，因此管理员必须自己确保集合的完备和正确性。

10.2.4.3 修改安全组

在安全组信息页面，先选中需要操作的安全组，再点击“修改”按钮，弹出安全组修改面板。该面板的操作都与新增安全组一致，因此这里就不在赘述，详情参考 10.2.4.2。

10.2.4.4 删除安全组

在过安全组信息页面，先选中需要删除的安全组信息，再点击“删除”按钮，弹出确认删除的面板。最后点击“确定”按钮执行删除操作。

10.2.4.5 安全组详细信息

在安全组信息页面，先选中需要查看的安全组，再点击“详细信息”按钮，弹出安全组详细信息面板，如图 10-27 所示。在该页面上不能修改安全组的信息。

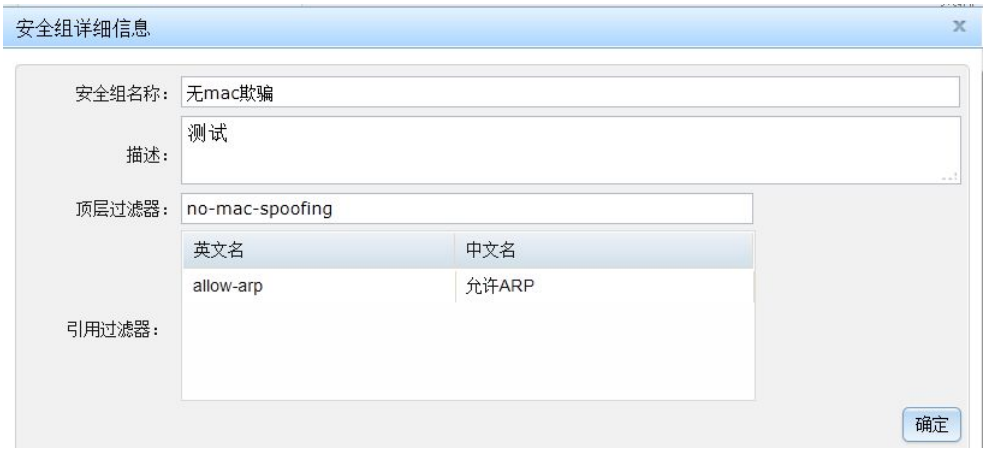


图 10- 27：安全组详细信息

10.3 项目管理员管理网络

项目管理员配置项目中虚拟机需要使用的网络。网络需要使用的子网和安全组由系统管理员配置好，然后分配给项目使用。由于桥接网络是和物理网络直接相连的，在物理网络中同一个 IP 子网只会出现在一个地方，因此给桥接网络使用的子网是不能被共享的。隔离网络和 NAT 网络的子网都不直接和物理网络相连，可以看做私有的网络，因此它使用的子网可以在多个其他地方再被使用。路由网络通过 IP 路由也要和物理网络互连的，因此其子网也不能共享。

安全组则是完全可以共享的，因此在配置虚拟机网络接口的时候，可以选择任意已经配置好的安全组。

如前所述，本系统将网络使用的子网和网络自身的配置分开。一个子网包含网络地址、网络掩码、网关地址、地址池（必须）、DHCP 服务、DNS 服务、tftp 服务和 bootp 服务。子网由系统管理员规划、配置并分配。项目管理员只需配置网络实现本身。

本系统将网络的信息分成以下几个部分：

1） 网络描述：包括网络的中名称、英文名称和描述信息。其中中文名称和描述都是为

了便于理解和记忆网络用项目管理员输入的。英文名称则要求是唯一的，以便在同一台宿主机上可以同时部署多个网络。

2） 网络基本信息：这是网络的基本设置，包括以下各项：

- 桥接器名称：部署虚拟网络使用的桥接器的名称，建议根据网络类型规划名称，以便区分桥接器是给那种网络使用的。
- 桥接器 MAC 地址：该项仅对隔离、路由和 NAT 网络有效，用于设定桥接器的 MAC 地址，但是建议不要配置，由系统自动生成。
- 绑定到宿主机：即将网络限定在某台宿主机上。这使得系统只在给定的宿主机上建这个虚拟网络。隔离、路由和 NAT 网络必须绑定宿主机。桥接网络则不需要。
- 启用 IPv6：该项对隔离、路由和 NAT 网络有效，即是否支持 IPv6。桥接网络该项没有意义。
- 启用 STP：STP 指生成树协议，该协议可应用于在网络中建立树形拓扑，消除网络中的环路。建议都选择。
- 转发网络接口：指定宿主机上用于构建虚拟网络的物理接口。隔离网络不能设置转发接口，NAT、路由、Linux 桥接、ovs 桥接网络可配置一个接口，用于指定和外部网络互连的宿主机接口。macvtap 可以配置多个物理接口，系统自动在多个接口间均衡流量。不过，NAT 和路由网络可以不指定转发网络接口，这时系统使用所有可以转发 IP 包的接口。
- 转发延迟时间：该项建议不要配置，当需要模拟路由或者交换设备的通信延迟时可以设定它。
- 使用的子网：虚拟网络使用的子网。在下拉框中选择。注意被其他桥接网络使用的子网不能再被选择。

3） 网络流量限制，网络流量限制仅对隔离、NAT 和路由网络有效。它用于限定进入和离开桥接器的总流量。

4） 静态路由配置：该部分仅对隔离、NAT 和路由网络有效。在建立虚拟机网络时，系统自动将静态路由配置加到宿主机的路由配置中，从而满足虚拟网络内通信的需求。

5） 端口组配置：该部分对所有网络都有效。端口组是一种配置端口通信限制的手段。将相同限制要求的端口作为一个组，并制定一个名称。在配置虚拟机接口时，为接口指定一个端口组就是指定一个通信限制。端口组配置中，流量限制是都可以配置。只有 OVS 网络的端口组还可以配置端口的 VLAN 信息。

6） NAT 网络转发配置：该部分配置仅对 NAT 网络有效，用于配置在虚拟网络和物理网络间转发数据包的规则。缺省允许和所有的地址和端口通信，设定地址范围或者端口范围后将只允许和设定范围内的主机和端口通信。

7） 网络 VLAN 配置：该部分仅对 OVS 桥接网络有效，只有 OVS 桥接网络支持 VLAN。

表 10-1 是各种网络可以配置部分的说明。

表 10-1：各种类型网络的配置情况

网络类型	描述信息	基本信息	流量限制	静态路由	端口组	NAT 转发配置	VLAN 配置
隔离网络	必须	必须	可选	可选	可选	无	无
NAT 网络	必须	必须	可选	可选	可选	可选	无
路由网络	必须	必须	可选	可选	可选	无	无
Linux 桥接网络	必须	必须	无	无	可选	无	无
OVS 桥接网络	必须	必须	无	无	可选	无	可选
Macvtap 桥接网络	必须	必须	无	无	可选	无	无

项目管理员的网络管理页面的内部结构如图 10-28 所示。

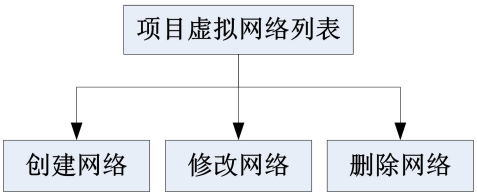


图 10- 28：网络管理内部结构图

10.3.1 网络信息

项目管理员的网络管理的首页是网络信息页面，如图 10-29。这个页面以列表的形式列出了当前项目管理员的所有网络。项目管理员可以根据网络名称来过滤数据。数据结果分页显示。列表顶端只有三个常用功能按钮：新增、修改、删除。


网络信息

新增 修改 删除

名称： 查询 重置

序号	中文名称	英文名称	桥接器名称	桥接器MAC地址	主机网络接口	网络类型	绑定主机名称	子网网关地址	子网名称
1	赵的linux桥接网络-有子网	zhao-linux-br-subnet	zhaobr2		eth1	Linux桥接网络		192.168.20.1	赵 192.168.20.0
2	赵的macvtap网络	zhao-macvtap			eth0,eth1	MACVTAP桥接网络			
3	赵的OVS网络-有子网	zhao-ovs-subnet	zhao-ovsbr2		eth1	OVS桥接网络		192.168.22.1	赵 -192.168.22.0
4	赵的路由网络	zhao-rout-net	zhao-route-net		eth1	路由网络	server188.txy.com	192.168.21.0	赵 192.168.21.0
5	赵的ovs网络-无子网	zhao-ovs-br	peng5		eth1	OVS桥接网络	server185.txy.com		
6	赵的NAT网络-有子网	zhao-nat-subnet	zhao-virbr0		eth0	NAT网络	189.txy.com	192.168.21.0	赵 192.168.21.0
7	赵的内部网络	zhao-innetnet	zhao-virbr0			隔绝网络	189.txy.com	10.1.0.1	内部服务子网 (10.1.0.0)
8	赵的linux桥接网络-无子网	zhao-linux-br	zhaobr0		eth1	Linux桥接网络			

第 1 - 8 个（共 8 项目数） 1

图 10- 29：网络信息列表

10.3.2 新增网络

本系统目前支持六种网络：隔绝网络、NAT 网络、路由网络、Linux 桥接网络、OVS 桥接网络和 macvtap 桥接网络。

其中，网络可能需要使用子网。隔绝网络和 NAT 网络使用的子网的网关不能设置为自动建立，并且可以重复使用子网；路由网络使用的子网的网关也不能设置为自动建立，并且不能与桥接网络和路由网络重复使用；macvtap 桥接网络不使用子网；其余的桥接网络（Linux 桥接和 OVS 桥接）不能与桥接网络和路由网络重复使用，它们都是独占式的。

10.3.2.1 新增隔绝网络

新增隔绝网络页面共有五个子面板：网络描述信息、网络基本信息、网络流量限制、静态路由配置、端口组配置。

(一) 网络描述信息，如图 10-30 所示。

1) 网络英文名称：在系统中的唯一名称，只允许字母、数字、中划线、下划线和点。
2) 网络中文名称：网络的中文名称，方便记录和辨认。
3) 网络描述：对当前网络的详细说明，仅用于方便用户对网络进行了解。

图 10- 30：新增隔绝网络之网络描述信息

(二) 网络基本信息，如图 10-31 所示。

1) 桥接器名称：定义将用于构建虚拟网络的桥接设备的名字。虚拟机将被连接到这个桥接设备，以便他们之间可以相互通信。建议桥接设备的名字以“vir”为前缀，但是“virbr0”保留给“default”虚拟网络。
2) 桥接器 MAC 地址：定义一个 MAC 地址，当网络创建时，该地址赋给桥接设备。最好不要设置 MAC 地址。如果需要，系统将自动产生一个随机的 MAC 地址。
3) 绑定到宿主机：指定当前网络需要绑定的宿主机。该类型的网络必须绑定宿主机。
4) 启用 IPv6：支持 Ipv6 地址。
5) 启用 STP：STP 指生成树协议，该协议可应用于在网络中建立树形拓扑，消除网络中的环路。
6) 转发延迟时间（毫秒）：桥接器转发延迟时间，建议不设置，默认值为 0。
7) 使用的子网：当前网络使用的子网。

8) 使用的子网详细信息：子面板中列出了选中子网的基本信息。

图 10- 31：新增隔绝网络之网络基本信息

(三) 网络流量限制，如图 10-32 所示。

流量限制指所有连接到网络的虚拟机接口的进出带宽总和。“平均”指桥接器流量的平均速率。“尖峰”指桥接器能发送的最大速率，输出时此值被忽略。“突发“指在尖峰速度时单个突发可以传输的最大 KB 数。

图 10- 32：新增隔绝网络之网络流量限制

(四) 静态路由配置，如图 10-33 所示

静态路由为虚拟主机提供路由信息，以便到达虚拟主机不能直接到达的网络，但是从宿

主机是可以直接到达的。例如，某个虚拟机配置了两个网络接口，分别连接两个虚拟网络。现在该虚拟机扮演路由角色，可在两个虚拟网络间路由 IP 包。则可能需要在宿主机上配置两条静态路由，以便将从一个虚拟机网络到达另一个虚拟网络的包转发给该虚拟机。新增面板（图 10-34）的各项含义如下：

图 10- 33：新增隔绝网络之静态路由配置

1) IP 版本：指配置的 IP 地址是使用 IPv4 还是 IPv6。
2) 网络地址：指要达到的目标主机或 IP 网段。
3) 网络掩码：网络地址的子网掩码前缀，使用数字表示，IPv6 最大支持 128，IPv4 最大支持 32，例如 IPv4 的一个网段 192.168.0.0，那么网络掩码则必须是 24，IPv6 的一个网段是 2001:DA6:0201:，那么网络掩码则必须是 48。
4) 网关地址：指下一跳路由入口 IP 地址。

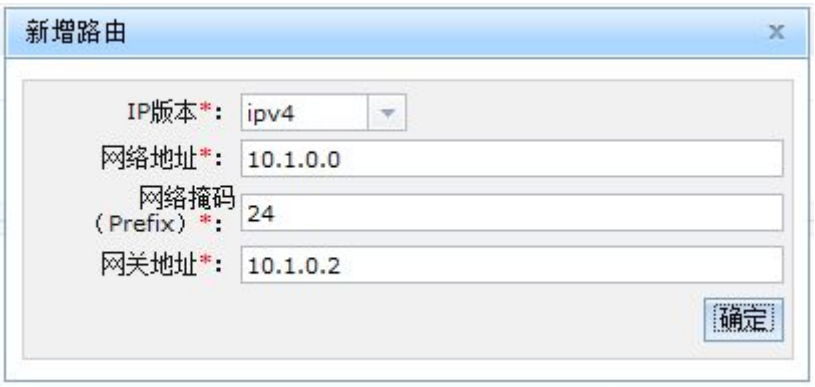


图 10- 34：新增隔绝网络之新增路由

(五) 端口组配置，如图 10-35 所示。

端口组配置给虚拟机的网络接口所使用。如果网络配置有端口组，而虚拟机网络接口没有指定端口组，则默认使用第一个端口组。如果没有配置端口组，则虚拟机网络接口不使用端口组。新增面板如图 10-36，名称指端口组在系统内的唯一标识，下面的流量控制指的是对网卡的流量限制。

图 10- 35：新增隔绝网络之端口组配置

图 10- 36：新增隔绝网络之新增端口组

10.3.2.2 新增 NAT 网络

新增 NAT 网络页面共有六个子面板：网络描述信息、网络基本信息、网络流量限制、静态路由配置、端口组配置、NAT 转发配置。除了 NAT 转发配置是 NAT 类型网络特有的外，其他面板的配置和隔绝网络是相同的。

(一) 网络描述信息，如图 10-37 所示。

1) 网络英文名称：在系统中的唯一名称，只允许字母、数字、中划线、下划线和点。

2) 网络中文名称：网络的中文名称，方便记录和辨认。

3) 网络描述：对当前网络的详细说明，仅用于方便用户对网络进行了解。

图 10- 37：新增 NAT 网络之网络描述信息

(二) 网络基本信息，如图 10-38 所示。

图 10- 38：新增 NAT 网络之网络基本信息

1) 桥接器名称：定义将用于构建虚拟网络的桥接设备的名字。虚拟机将被连接到这个桥接设备，以便他们之间可以相互通信。桥接设备不直接连接到物理网络，但由宿主机按照 NAT 方式在虚拟网络和物理网络间转发 IP 包。建议桥接设备的名字以“vir”为前缀，但是“virbr0”保留给“default”虚拟网络。
2) 桥接器 MAC 地址：定义一个 MAC 地址，当网络创建时，该地址赋给桥接设备。最好不要设置 MAC 地址。如果需要，系统将自动产生一个随机的 MAC 地址。
3) 绑定到宿主机：指定当前网络需要绑定的宿主机。该类型的网络必须绑定宿主机。
4) 启用 IPv6：支持 Ipv6 地址。
5) 启用 STP：STP 指生成树协议，该协议可应用于在网络中建立树形拓扑，消除网络中的环路。
6) 转发网络接口：将虚拟机的 IP 地址转换为宿主机的公共 IP 地址后使用指定的设备转发。
7) 转发延迟时间（毫秒）：桥接器转发延迟时间，一般不设置，默认值为 0。
8) 使用的子网：当前网络使用的子网。

9) 使用的子网详细信息：子面板中列出了选中子网的基本信息。

(三) 网络流量限制，如图 10-39 所示。

流量限制指所有连接到网络的虚拟机接口的进出带宽总和。平均指桥接器流量的平均速率。尖峰指桥接器能发送的最大速率，输出时此值被忽略。突发指在尖峰速度时单个突发可以传输的最大 KB 字节数。

图 10- 39：新增 NAT 网络之网络流量限制

(四) 静态路由配置，如图 10-40 所示。

静态路由为虚拟主机提供路由信息，以便到达虚拟主机不能直接到达，但是从宿主机是可以直接到达的网络。新增面板（图 10-41）的各项含义如下：

图 10- 40：新增 NAT 网络之静态路由配置

1) IP 版本：指配置的 IP 地址是使用 IPv4 还是 IPv6。
2) 网络地址：目标主机或 IP 网段。
3) 网络掩码：网络地址的子网掩码前缀，使用数字表示，IPv6 最大支持 128，IPv4 最大支持 32，例如 IPv4 的一个网段 192.168.0.0，那么网络掩码则必须是 24，IPv6 的一个网段是 2001:DA6:0201:，那么网络掩码则必须是 48。
4) 网关地址：指下一跳路由入口 IP 地址。

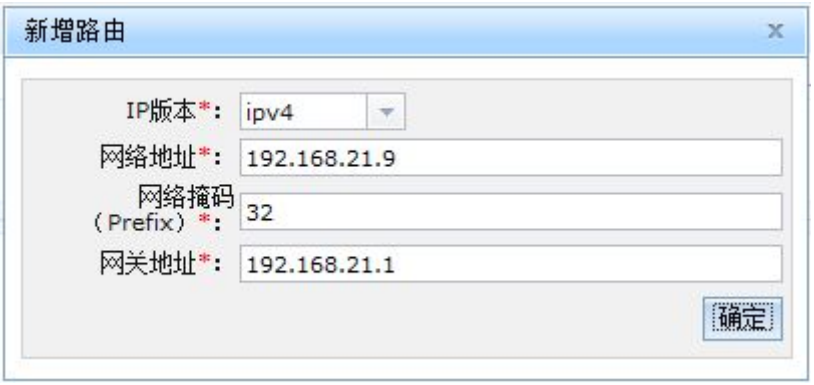


图 10- 41：新增 NAT 网络之新增路由

(五) 端口组配置，如图 10-42 所示。

端口组配置给虚拟机的网络接口所使用。如果网络配置有端口组，而虚拟机网络接口没有指定端口组，则默认使用第一个端口组。如果没有配置端口组，则虚拟机网络接口不使用端口组。新增面板如图 10-43，名称指端口组在系统内的唯一标识，下面的流量控制是对虚拟机网络接口的流量限制。

图 10- 42：新增 NAT 网络之端口组配置

图 10- 43：新增 NAT 网络之新增端口组

(六) NAT 网络转发配置配置，如图 10-44 所示。

连接到这个网络的虚拟机和物理网络间通过宿主机的 IP 协议栈，在将虚拟机的 IP 地址转换为宿主机的公共 IP 地址后转发。这使得运行在只有一个公共 IP 地址的宿主机上的多个虚拟机，都能访问物理网络。由于 IPv6 没有 NAT，如果网络中有一些 IPv6 地址，则 IPv6 通信将使用普通路由转发。防火墙规则将允许外出连接通过任何网络设备，无论是 ethernet、wireless、dialup 或者 VPN。如果在网络基本信息里设置了网络转发接口，防火墙规则将仅使用指定的设备转发，从其他网络设备进入的连接是阻止的。同一虚拟网络上的虚拟机之间的所有连接，包括宿主机和虚拟机之间的，都是不限制也不进行 NAT 的。

图 10- 44：新增 NAT 网络之 NAT 网络转发配置

新增地址范围如图 10-45 所示。配置允许转发的公共 IPv4 地址范围，即允许访问虚拟网络的公共 IP 地址范围。如果配置的是单个地址，则可以把起始地址和结束地址设为相同

值。如果没有设定任何值，缺省是允许所有的地址。

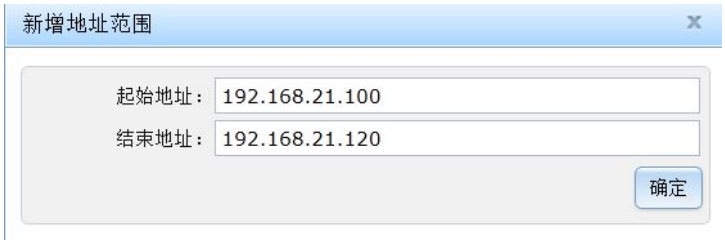


图 10- 45：新增 NAT 网络之新增网络转发地址范围

新增端口范围如图 10-46，配置允许公共 IP 允许连接的端口范围。如果没有设定任何值，缺省是允许所有的端口。

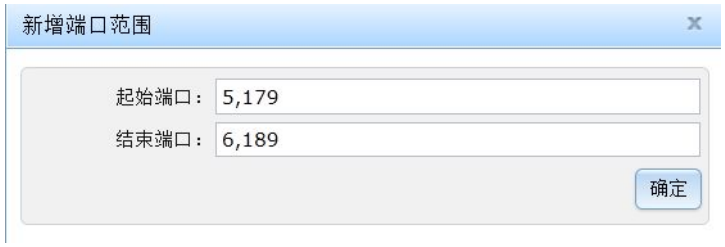


图 10- 46：新增 NAT 网络之新增网络转发端口范围

10.3.2.3 新增路由网络

新增路由网络页面共有五个子面板：网络描述信息、网络基本信息、网络流量限制、静态路由配置、端口组配置。路由网络的配置和隔绝网络的配置很相似，唯一不同的是路由网络可以配置转发网络接口，让虚拟网络和物理网络可以通过路由转发方式互联。

(一) 网络描述信息，如图 10-47 所示。

1) 网络英文名称：在系统中的唯一名称，只允许字母、数字、中划线、下划线和点。

2) 网络中文名称：网络的中文名称，方便记录和辨认。

3) 网络描述：对当前网络的详细说明，仅用于方便用户对网络进行了解。

图 10- 47：新增路由网络之网络描述信息

(二) 网络基本信息，如图 10-48 所示。

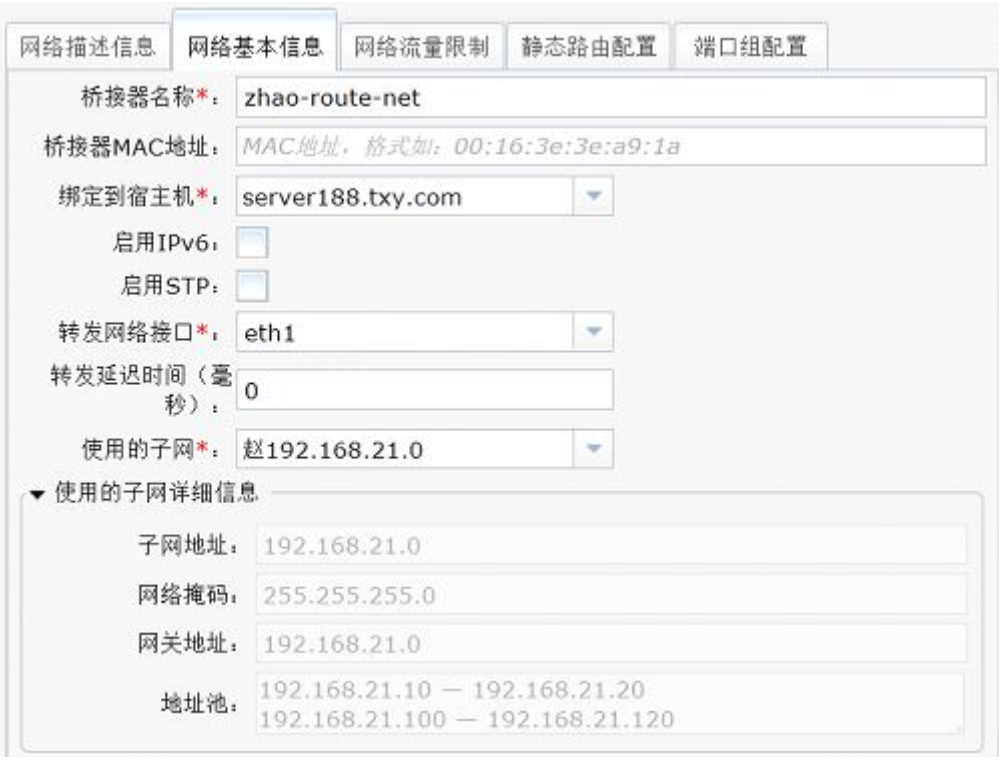


图 10- 48：新增路由网络之网络基本信息

1) 桥接器名称：定义将用于构建虚拟网络的桥接设备的名字。虚拟机将被连接到这个桥接设备，以便他们之间可以相互通信。建议桥接设备的名字以“vir”为前缀，但是“virbr0”保留给“default”虚拟网络。
2) 桥接器 MAC 地址：定义一个 MAC 地址，当网络创建时，该地址赋给桥接设备。最好不要设置 MAC 地址。如果需要，系统将自动产生一个随机的 MAC 地址。
3) 绑定到宿主机：指定当前网络需要绑定的宿主机。该类型网络必须绑定宿主机。
4) 启用 IPv6：支持 Ipv6 地址。
5) 启用 STP：STP 指生成树协议，该协议可应用于在网络中建立树形拓扑，消除网络中的环路。
6) 转发网络接口：客户机网络通信将通过宿主机的 IP 协议栈转发到物理网络，不使用 NAT，转发网络接口即数据包转发出去的接口。虚拟机进出的会话是不受限制的（可通过在虚拟机的网络接口上配置 nwfilter 规则来限制进入虚拟机的通信）。
7) 转发延迟时间（毫秒）：桥接器转发延迟时间，一般不设置，默认值为 0。
8) 使用的子网：当前网络使用的子网。
9) 使用的子网详细信息：子面板中列出了选中子网的基本信息。

(三) 网络流量限制，如图 10-49 所示。

流量限制指所有连接到网络的虚拟机接口的进出带宽总和。平均指桥接器流量的平均速率。尖峰指桥接器能发送的最大速率，输出时此值被忽略。突发指在尖峰速度时单个突发可以传输的最大 KB 字节数。

图 10- 49：新增路由网络之网络流量限制

(四) 静态路由配置，如图 10-50 所示。

静态路由为虚拟机提供路由信息，以便到达虚拟机不能直接到达，但是从宿主机是可以直接到达的网络。新增面板（图 10-51）的各项含义如下：

1) IP 版本：指配置的 IP 地址是使用 IPv4 还是 IPv6。
2) 网络地址：访问的目标主机或 IP 网段。
3) 网络掩码：网络地址的子网掩码前缀，使用数字表示，IPv6 最大支持 128，IPv4 最大支持 32，例如 IPv4 的一个网段 192.168.0.0，那么网络掩码则必须是 24，IPv6 的一个网段是 2001:DA6:0201:，那么网络掩码则必须是 48。
4) 网关地址：指下一跳路由入口 IP 地址。

图 10- 50：新增路由网络之静态路由配置

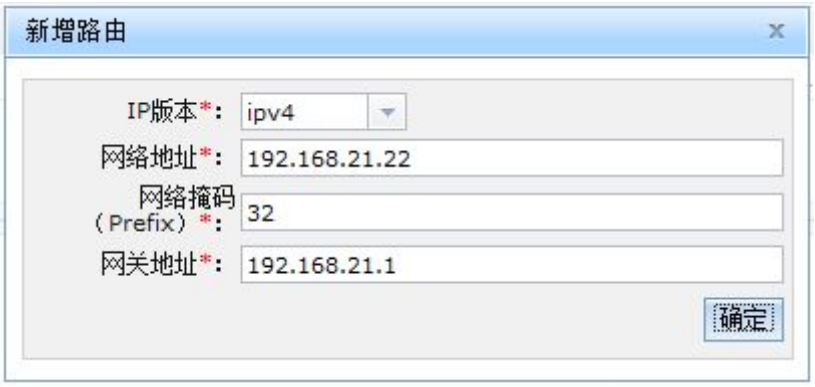


图 10- 51：新增路由网络之新增路由

(五) 端口组配置，如图 10-52 所示。

端口组配置给虚拟机的网络接口所使用。如果网络配置有端口组，而虚拟机网络接口没有指定端口组，则默认使用第一个端口组。如果没有配置端口组，则虚拟机网络接口不使用端口组。新增面板如图 10-53 所示。名称指端口组在系统内的唯一标识，下面的流量控制指的是对网卡的流量限制。

图 10- 52：新增路由网络之端口组配置

图 10- 53：新增路由网络之新增端口组

10.3.2.4 新增 Linux 桥接网络

桥接网络通常和物理网络直接互联（在链路层互通），因此 IP 子网的服务大多由物理路由器完成。这样，就无需配置静态路由和 NAT 转发。桥接网络也无法控制网络的整体流量，因此也没有网络流量配置。新增 Linux 桥接网络页面共有三个子面板：网络描述信息、网络基本信息、端口组配置。

(一) 网络描述信息，如图 10-54 所示。

1) 网络英文名称：在系统中的唯一名称，只允许字母、数字、中划线、下划线和点。

2) 网络中文名称：网络的中文名称，方便记录和辨认。

3) 网络描述：对当前网络的详细说明，仅用于方便用户对网络进行了解。

图 10- 54：新增 Linux 桥接网络之网络描述信息

(二) 网络基本信息，如图 10-55 所示。

1) 桥接器名称：定义将用于构建虚拟网络的桥接设备的名字。虚拟机将被连接到这个桥接设备，以便他们之间可以相互通信。指定的转发网络接口作为桥接的端口，从而使虚拟网络连接到物理 LAN。建议桥接设备的名字以“linux-”为前缀，如 linux-br0。

2) 绑定到宿主机：指定当前网络需要绑定的宿主机。桥接网络可以跨宿主机部署，因

此建议不要绑定。

3) 启用 IPv6：支持 Ipv6 地址。

4) 启用 STP：STP 指生成树协议，该协议可应用于在网络中建立树形拓扑，消除网络中的环路。

5) 转发网络接口：宿主机上配置给桥接器使用的物理接口。此接口不能再被宿主机使用。

6) 转发延迟时间（毫秒）：桥接器转发延迟时间，一般不设置，默认值为 0。

7) 使用的子网：当前网络使用的子网。

8) 使用的子网详细信息：子面板中列出了选中子网的基本信息。

图 10- 55：新增 Linux 桥接网络之网络基本信息

(三) 端口组配置，如图 10-56 所示。

端口组是给虚拟机的网络接口配置的，默认值为第一组数据。如果没有配置端口组，则在配置虚拟机的网络接口时不能使用端口组。新增面板如图 10-57，名称指端口组在虚拟网络内的唯一标识。下面的流量控制指的是对网卡的流量限制。

图 10- 56：新增 Linux 桥接网络之端口组配置

图 10- 57：新增 Linux 桥接网络之新增端口组

10.3.2.5 新增 OVS 桥接网络

OVS 桥接网络是桥接网络的一种。和 Linux 桥接网络相比，OVS 桥接网络不仅具有 Linux 桥接网络的功能，还支持 VLAN，因此 OVS 网络有单独的网络 VLAN 配置面板。新增 OVS 桥接网络页面共有四个子面板：网络描述信息、网络基本信息、端口组配置、网络 VLAN 配置。

(一) 网络描述信息，如图 10-58 所示。

1) 网络英文名称：在系统中的唯一名称，只允许字母、数字、中划线、下划线和点。

2) 网络中文名称：网络的中文名称，方便记录和辨认。

3) 网络描述：对当前网络的详细说明，仅用于方便用户对网络进行了解。

图 10- 58：新增 OVS 桥接网络之网络描述信息

(二) 网络基本信息，如图 10-59 所示。

1) 桥接器名称：定义将用于构建虚拟网络的桥接设备的名字。虚拟机的网络接口将被连接到这个桥接设备，以便他们之间可以相互通信。指定的转发网络接口也连接到桥接设备，以便将虚拟网络连接到物理 LAN。建议桥接设备的名字以“ovs-”为前缀，例如 ovs-br0。

2) 绑定到宿主机：指定当前网络需要绑定的宿主机。OVS 虚拟网络可以跨宿主机部署，建议不绑定宿主机。
3) 启用 IPv6：支持 Ipv6 地址。
4) 启用 STP：STP 指生成树协议，该协议可应用于在网络中建立树形拓扑，消除网络中的环路。
5) 转发网络接口：配置给桥接器使用的宿主机的物理接口。此设备不能再被宿主机使用。
6) 转发延迟时间（毫秒）：桥接器转发延迟时间，一般不设置，默认值为 0。
7) 使用的子网：当前网络使用的子网。
8) 使用的子网详细信息：子面板中列出了选中子网的基本信息。

图 10- 59：新增 OVS 桥接网络之**网络基本信息**

(三) 端口组配置，如图 10-60：

图 10- 60：新增 OVS 桥接网络之端口组配置

端口组是给虚拟机的网络接口配置的，默认值为第一组数据。如果没有配置端口组，则在配置虚拟机的网络接口时不能使用端口组。新增面板如图 10-61，名称指端口组在系统内的唯一标识，下面的流量控制指的是对虚拟机网络接口的流量限制。vlan 配置指虚拟机网络

接口连接的对应桥接器端口的 vlan 配置，详细配置可以参考下一项的网络 VLAN 配置。

图 10- 61：**新增**新增 OVS 桥接网络之新增端口组

(四) 网络 VLAN 配置，如图 10-62 所示。

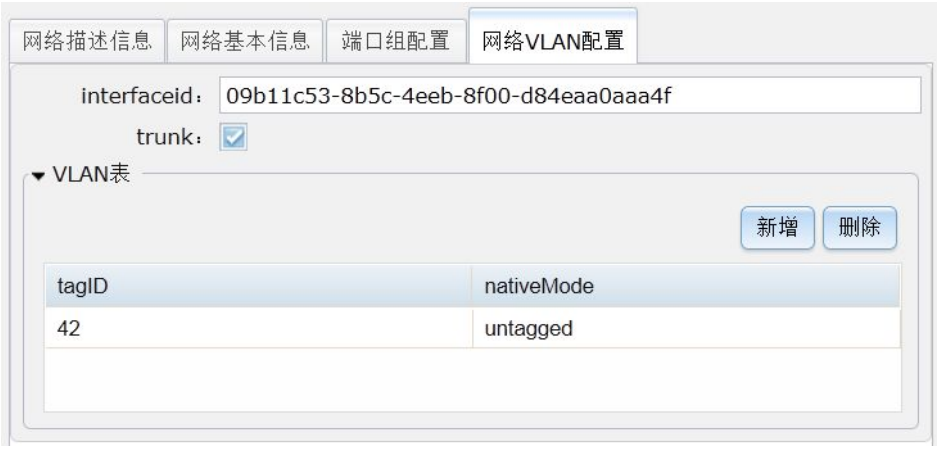


图 10- 62：新增 OVS 桥接网络之网络 VLAN 配置

网络的 VLAN 是给所有连接该网络的虚拟机接口的默认 VLAN。如果在端口组中配置了 VLAN，则可被端口组配置覆盖。而端口组的 VLAN 配置又可被直接在虚拟机接口中的 VLAN 配置覆盖。给 OVS 网络配置 VLAN 要注意：连接宿主机的交换机也要支持 VLAN 才能使跨宿主机部署的 OVS 虚拟网络互连，否则可能无法跨宿主机通信。

interfaceid：指虚拟机网络接口 uuid。一般不指定，在第一次定义这个接口时，将随机生成一个 uuid。

trunk：虚拟局域网中继技术，让连接在不同交换上的相同 VLAN 中的主机互通。即 trunk 类型的端口可以允许多个 VLAN 通过，可以接收和发送多个 VLAN 的报文。

VLAN 表：允许通过该接口的 VLAN。新增面板如图 10-63。tagID 是 VLAN 的唯一标识，是一个十进制数字。nativeMode 是 VLAN 在该端口上的状态，有两个有效值：tagged、untagged，还有一个 none 表示不配置此值。一个接口只可以有一个 untagged 模式的 VALN。

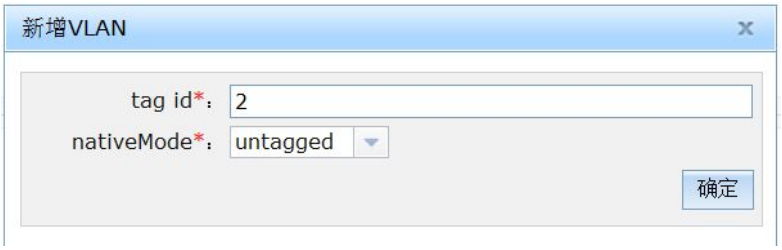


图 10- 63：新增 OVS 桥接网络之新增 VLAN 配置

10.3.2.6 新增 MACVTAP 桥接网络

Macvtap 桥接网络既不支持 VLAN，也不支持端口组，只能配置网络描述信息和网络基本信息。采用 Macvtap 桥接网络时，所有的网络服务，包括网关、DHCP 服务等都必须有系统外部提供，即由物理路由器等提供。本系统只负责将虚拟机网络接口的数据包桥接到物理网络上。

(一) 网络描述信息，如图 10-64 所示。

1) 网络英文名称：在系统中的唯一名称，只允许字母、数字、中划线、下划线和点。
2) 网络中文名称：网络的中文名称，方便记录和辨认。
3) 网络描述：对当前网络的详细说明，仅用于方便用户对网络进行了解。

图 10- 64：新增 MACVTAP 桥接网络之网络描述信息

(二) 网络基本信息，如图 10-65 所示。

1) 绑定到宿主机：指定当前网络需要绑定的宿主机。如果使用了特定宿主机的物理接口，或者用于桥接的物理接口在其他宿主机上没有，则必须绑定到具体的宿主机。否则可以不绑定。
2) 启用 IPv6：支持 IPv6 地址。
3) 转发网络接口表：宿主机的网络接口被用于直接连接。在链路层，虚拟机的接口就像直接连接到物理接口。可以同时指定多个转发网络接口。

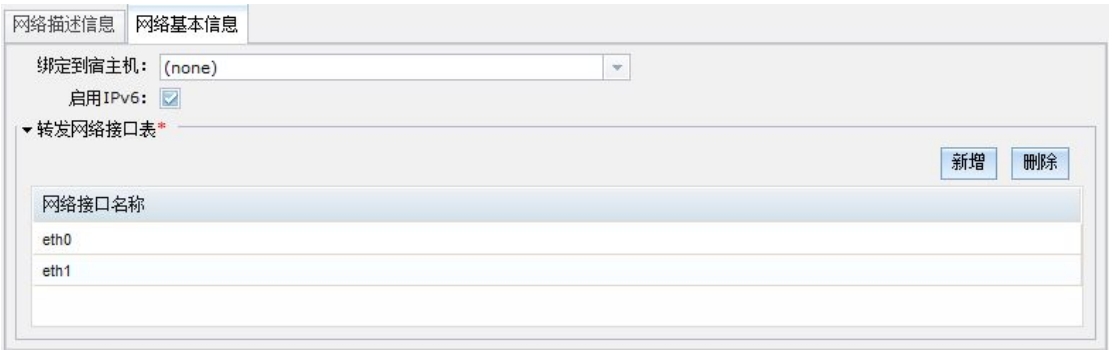


图 10- 65：新增 MACVTAP 桥接网络之网络基本信息

10.3.3 修改网络

在网络信息页面，先选中需要操作的网络，再点击“修改”按钮，页面跳转至网络配置页面，界面与操作都与新增网络一致，因此，这里就不再赘述，详情参考 10.3.2。

10.3.4 删除网络

在网络信息页面，先选中需要删除的网络，再点击“删除”按钮，弹出确认删除的页面，最后点击“确定”按钮执行删除操作。*注意，如果此网络正在被虚拟机使用，则删除操作将会执行失败。*

11. 运行虚拟机

本小节介绍虚拟机实例和虚拟机，它们属于“项目管理员”的管理范畴。

11.1 虚拟机运行环境概述

虚拟机管理是项目管理的核心。图 11-1 是虚拟机管理页面组内部结构。要注意的是：一个项目有许多虚拟机实例，每个虚拟机实例可以有一个或者多个虚拟机存在于系统中。图中功能的组织正是体现了这种结构。

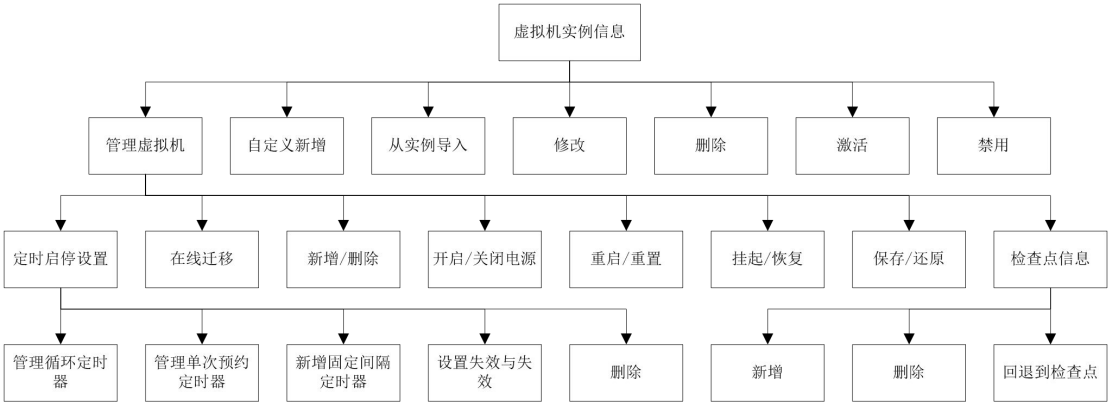


图 11- 1：虚拟机实例管理页面组织

本部分的首页面是“虚拟机实例信息”页面。这个页面以列表的方式显示项目中的所有虚拟机实例。列表字段包括：中文名称、英文名称、绑定宿主机、VCPU 数、内存大小、管理状态，虚拟机数量和创建时间。另外一个页面是“项目虚拟机信息”，只提供虚拟机的管理，但能够管理项目内的所有虚拟机。两者管理虚拟机的内容是大体一致的，因此，以下只介绍“虚拟机实例信息”页面。

11.2 为虚拟机准备环境

数字有机体虚拟机系统运行需要一定的环境。具体包括硬件和虚拟硬件两方面。硬件主要包括服务器（包括了真实的硬盘、真实的内存、真实的网卡）和网络（包括真实的网线、路由器、交换机等），它们是所有软件的载体；虚拟硬件主要包括虚拟服务器、虚拟硬盘、虚拟网络（或网卡）、虚拟内存等。

硬件设备是必不可少的，可以直接使用；虚拟硬件是构建于硬件之上的，无法直接使用，需要通过软件来配置和制作之后才能被使用。各种种类的虚拟硬件功能各不相同，因此虚拟各种硬件时的配置也有很多不同的地方，有的虚拟硬件配置复杂，有的虚拟硬件配置简单。硬盘和网络的虚拟化规格多种多样，并且参数设置较多，配置起来较为复杂。CPU 和内存参数较少，配置起来较简单。此外，虚拟化的硬盘设备是需要持久化特性的，虚拟硬盘设备应该和真实的硬盘设备一样，能够长时间保存文件。

针对硬件的虚拟化，数字有机体虚拟机系统开发了虚拟网络、虚拟硬盘（包括池和卷）

的管理页面。为了给虚拟服务器安装操作系统，还虚拟了启动驱动器（如 VCD 和 DVD 等）。

按照前面的描述，用户在使用虚拟机前，应该准备以下环境：

1） 服务器（包括硬盘、CPU、内存、网卡等硬件）；

2） 能使服务器相互联通的网络（包括网线、路由器和交换机等硬件）；

3） 制作虚拟的网络设备（如虚拟网络）；

4） 制作虚拟的硬盘（如卷）；

5） 如果需要安装软件（如操作系统），还需要制作虚拟的安装启动光盘（如 cd）。

虚拟 CPU 和虚拟内存在配置虚拟机时配置，无需提前准备。但是需要保证有足够多的资源来完成虚拟 CPU 和虚拟内存的运行。如果准备好了，就能够开始配置虚拟机了。

11.3 虚拟机实例管理

如果需要，我们可以直接配置一台虚拟机，并投入使用。但是，很多时候，特别是在大规模，大数量地使用虚拟机的时候，不少虚拟机的配置可能大同小异。直接配置虚拟机不仅容易出错，并且工作量也很大。针对这种情况，我们可以先新建一个虚拟机实例，然后把这个实例保存起来，当需要新建与这个实例大同小异的虚拟机的时候，我们只需要使用“从实例导入”把这个实例拷贝过来，再基于它进行修改就可以了。

虚拟机的实例管理主要包括新建、删除和修改，如图 11-2 所示。其中的新建可通过引用已有的实例，然后以修改的方式来实现；也可以完全自定义新建。其次还包括“激活”与“禁用”，被禁用的虚拟机实例是不能使用的，默认新建的虚拟机实例总是处于激活状态。

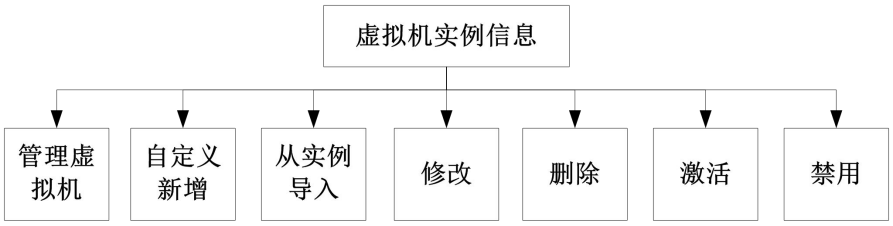


图 11- 2：实例管理页面组织

用户在定义实例的时候，首先应该确定需要虚拟什么型号的、32 位还是 64 位的主机，其次还要考虑分配多大的磁盘和采用何种网络。如果需要在虚拟机上安装新操作系统，可能还应该配置虚拟的光驱。

在导航栏点击“虚拟机管理”条目即进入虚拟机实例管理的用户界面。这个界面如图 11-3 所示。

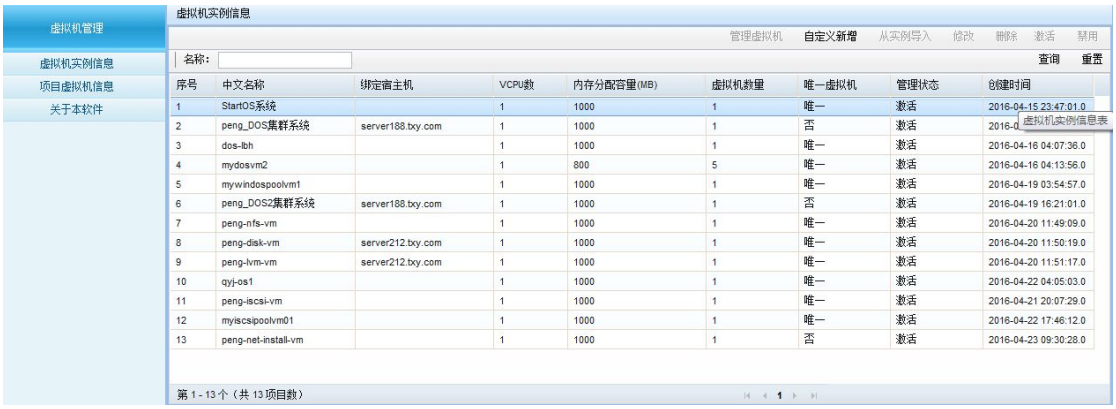


图 11- 3：虚拟机实例管理界面

新定义一个虚拟机的实例，可以使用“自定义新增”按钮；也可以选中一个实例，然后使用“从实例导入”按钮来实现。新增虚拟机实例共分为“基本配置”、“磁盘设备配置”、“图形界面配置”、“网络接口配置”和“用户自定义配置”五个方面的配置。正确配置各个方面后，给虚拟机实例填写一个中文名称之后，可以点击“保存”按钮进行提交。 以下是虚拟机实例配置的步骤：

第一步，配置虚拟机的实例基本信息。一般情况下可以只简单地修改 CPU 和内存的配置，其余均可采用缺省的配置，如下图所示。

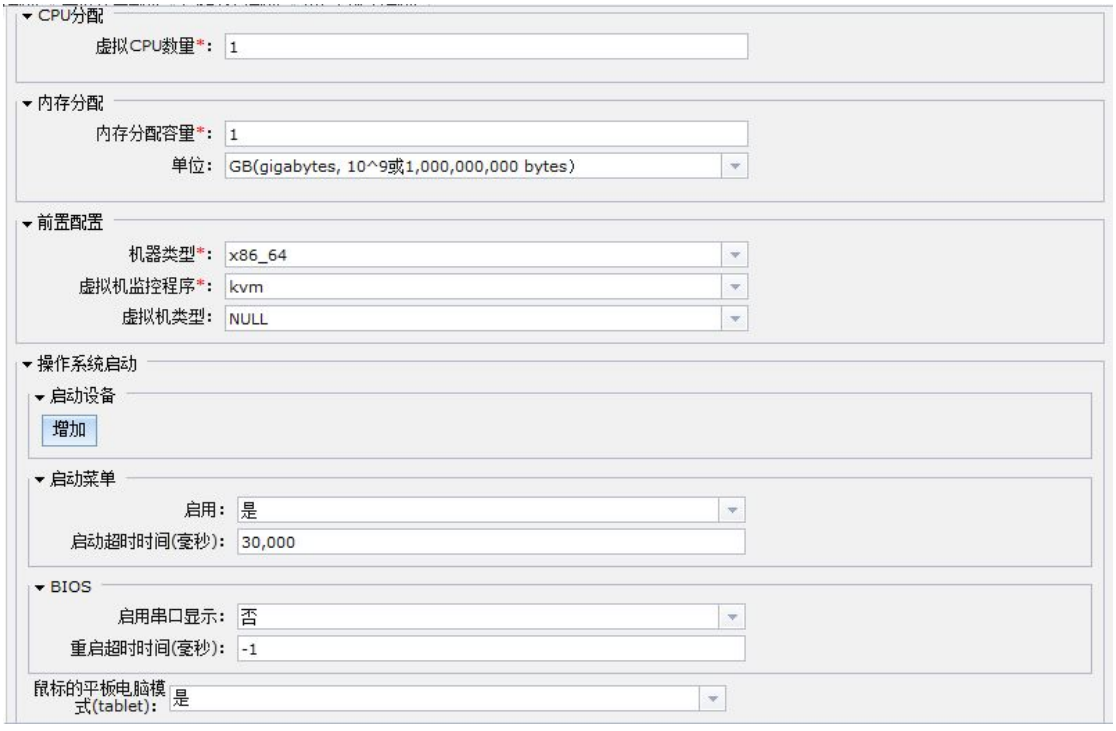


图 11- 4：基本配置

“基本配置”的各个配置选项的参数的填写请参考下表“实例的基本配置”。

表格 11-1 实例的基本配置

编号	配置参数	选项说明	功能描述
1	机器类型	目前支持“i686”和“x86_64”	虚拟主机的类型，“i686”表示 32 位主机；“x86_64”表示 64 位主机
2	虚拟机监控程序	目前支持“kvm”和“qemu”	在 Linux 系统下 KVM 的性能更好，而且 KVM 完全融入到 Linux 内核中。QEMU 模拟器则整合 XEN 和 KVM，提供完善的虚拟机运行功能
3	虚拟机类型	厂家推出的主机类型	虚拟真实服务器的型号
4	启动设备	目前支持“光盘”、“硬盘”、“软盘”和“网络”	这里描述了服务器启动的设备，例如常常使用 cd 安装系统，就需要设置“光盘”启动；这里可以选择多个启动驱动设备，位置越靠前越优先使用
5	是否启用“启动菜单”	开关量，选择“是”表示开，“否”表示关	选择“是”，虚拟机启动时会提示选择启动设备，但是超时会自动选择；选择“否”则不会提示选择
6	启动超时时间	单位为毫秒的正整数；-1 表示不超时。	提示选择启动设备的最长时间
7	是否启用串口显示	开关量，选择“是”表示开，“否”表示关	是否使用“串口连接显示设备”，选择“是”表示使用串口；选择“否”则不使用
8	重启超时时间	单位为毫秒的正整数；-1 表示不超时	系统启动超时时间，如果设置为正数，则在该时间内虚拟机未能成功启动将会自动重启
9	虚拟 CPU 数量	1 到 255 的非负数	表示虚拟的 CPU 的数量
10	内存分配容量	单位可选的容量	虚拟内存不得大于服务器的真实内存，一般地，需要保留至少 512MB 的内存给真实服务器上的系统，即：可用的虚拟内存最大值为真实的内存值减去 512MB
11	鼠标的平板电脑模式(tablet)	开关量，选择“是”表示开，“否”表示关	选择“是”时，Windows 操作系统的鼠标箭头和鼠标原点完全重合，方便使用，虚拟 Windows 和 FreeBSD 操作系统时最好选择“开”

第二步，配置虚拟机的磁盘。一个虚拟机最多支持四块磁盘，磁盘的类型可以分为

cdrom、disk 和 volume，每种磁盘的配置均有不同，如下图所示。

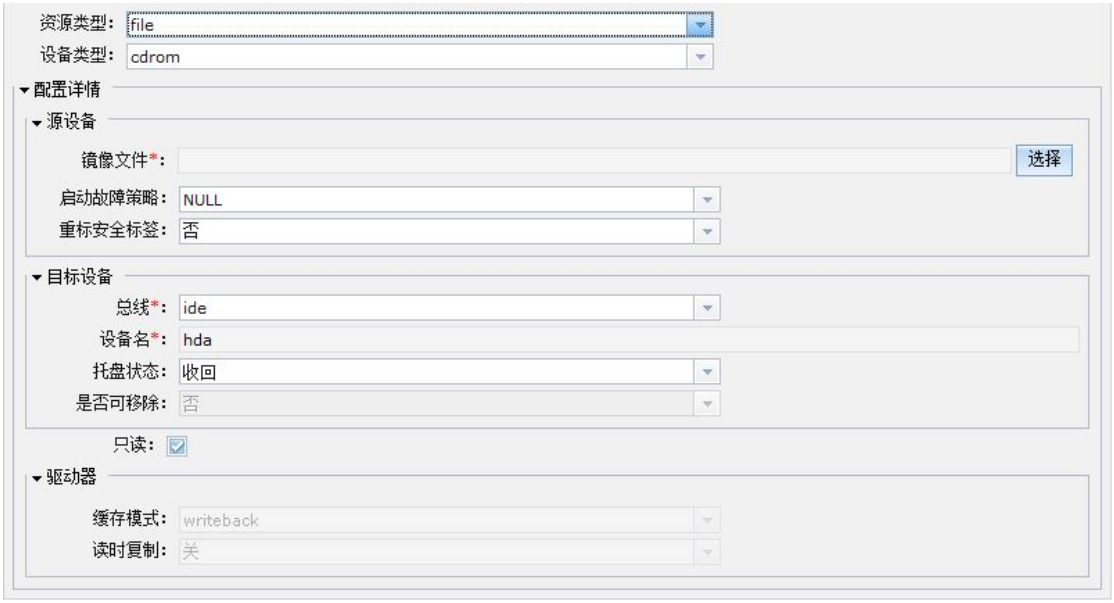


图 11- 5：磁盘设备 cdrom 配置

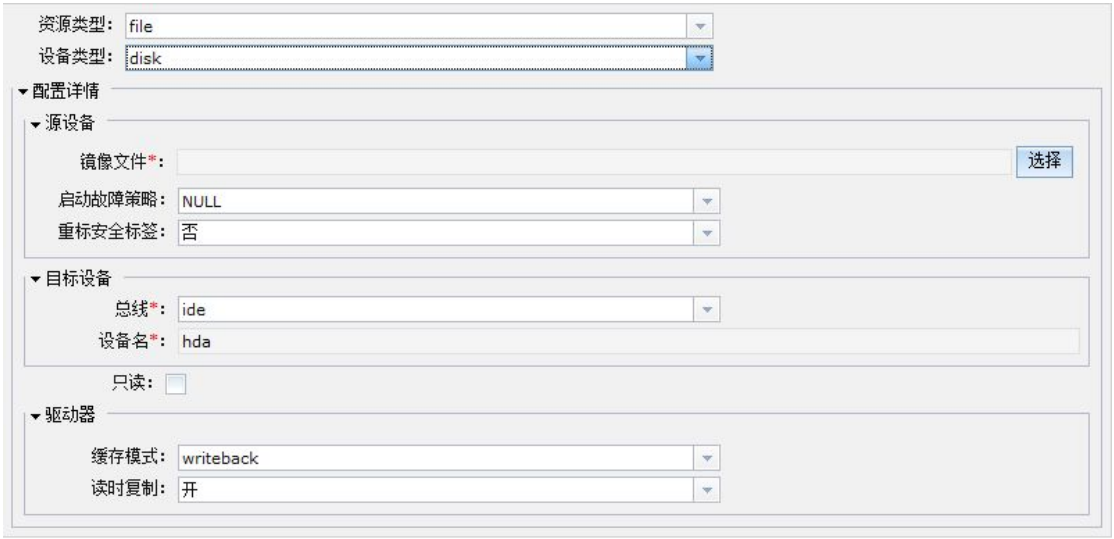


图 11- 6：磁盘设备 disk 配置

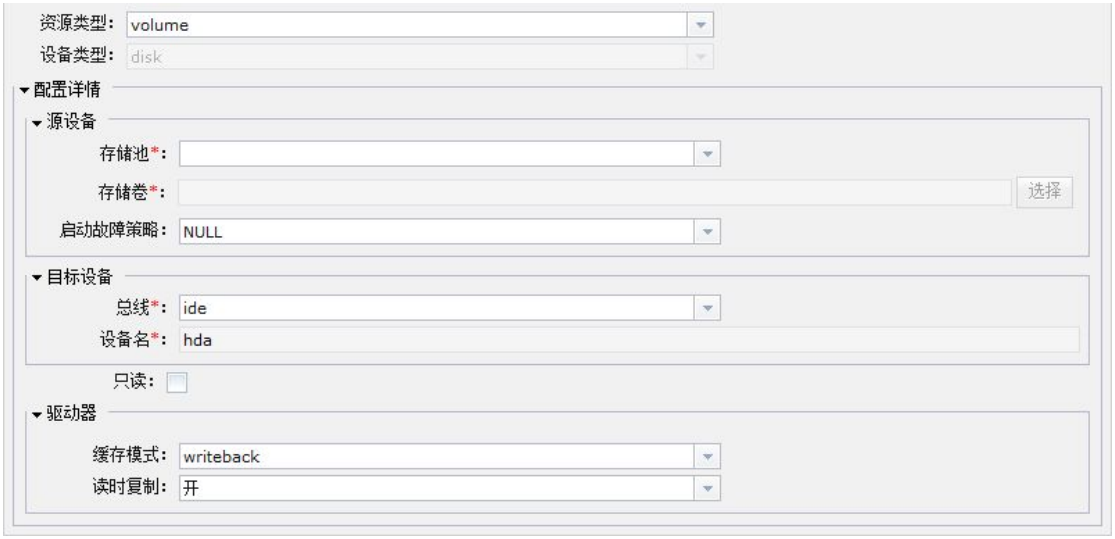


图 11- 7：磁盘设备 volume 配置

“cdrom”、“disk”和“volume”配置选项的参数的填写请参考表“实例的磁盘设备配置”。

表格 11-2 实例的磁盘设备配置

编号	配置参数	选项说明	功能描述
1	资源类型	目前支持“file”和“volume”	“file”的设备类型包括“disk”和“cdrom”；而“volume”的设备类型仅仅为“disk”
2	设备类型	目前支持“disk”和“cdrom”	此选项取决于“资源类型”
3	驱动器的缓存模式	缓存模式，目前支持的有：Default、None、Directsync、Unsafe、Writeback、Writethrough	属性 cache 控制缓存机制，可能的值有：default、none、writethrough、writeback、directsync（类似 writethrough，但它跳过主机页面缓存）、unsafe（主机可能缓存所有磁盘 io，并忽略从 guest 来的同步请求）
4	驱动器的读时复制	开关量，选择“开”表示开，“关”表示关	属性 copy_on_read 控制是否将读取的备份文件复制到镜像文件，它的值可以是“开”或“关”。copy_on_read 避免多次访问相同的备份文件分区，对在一个网速慢的网络中备份文件有用。copy_on_read 的默认值是“关”。 如果该磁盘的读时复制为“开”时，制作包含该磁盘的检查点将不能成功。
5	目标设备的总线	属性 bus 指定了被模拟的磁盘类型。	典型值有：ide、scsi、virtio、usb

6	目标设备的设备名	名称自动生成，名如 sda、sdb……，或者 hda、hdb……	逻辑设备的名字。不过，指定的名称并不能保证映射为虚拟机操作系统上的设备名
7	目标设备的托盘状态	可选择的有“收回”和“弹出”	默认是“收回”状态，表示光驱的状态
8	目标设备的是否可移除	选择“是”表示可移除，否则不可移除	总线选择“usb”时起效
9	只读	复选框，被选中时具备此特性	资源类型为 cdrom 的设备默认具有此特性，某些总线的磁盘设备不能是只读的。
10	源设备的存储池	选择存储池	指定磁盘源所有的存储池的名称。
11	源设备的存储卷	选择存储池中的卷	指定当作磁盘源的存储卷的名称
12	源设备的启动故障策略	目前支持以下选项：Mandatory、requistite、optional	对于 type 为 file 或 volume，device 为 cdrom 或 floppy 的磁盘，可以配置源文件不能访问时的处理策略：mandatory：因某些原因找不到则报失败（默认），requistite：在启动时找不到报失败，迁移、修复、恢复时找不到则卸载那些磁盘，optional：任何开始尝试时找不到就卸载那些磁盘
13	源设备的镜像文件	镜像文件的路径名，可选择。	在资源类型为 file 时有效。
14	源设备的重标安全标签	开关量，选择“是”表示开，“否”表示关	“是”表示 selinux 的安全标记需要重新标记

第三步，配置虚拟机的远程桌面。目前只支持 vnc 和 spice，如果使用 vnc，采用缺省配置即可。

图 11- 8：图形界面 vnc 配置

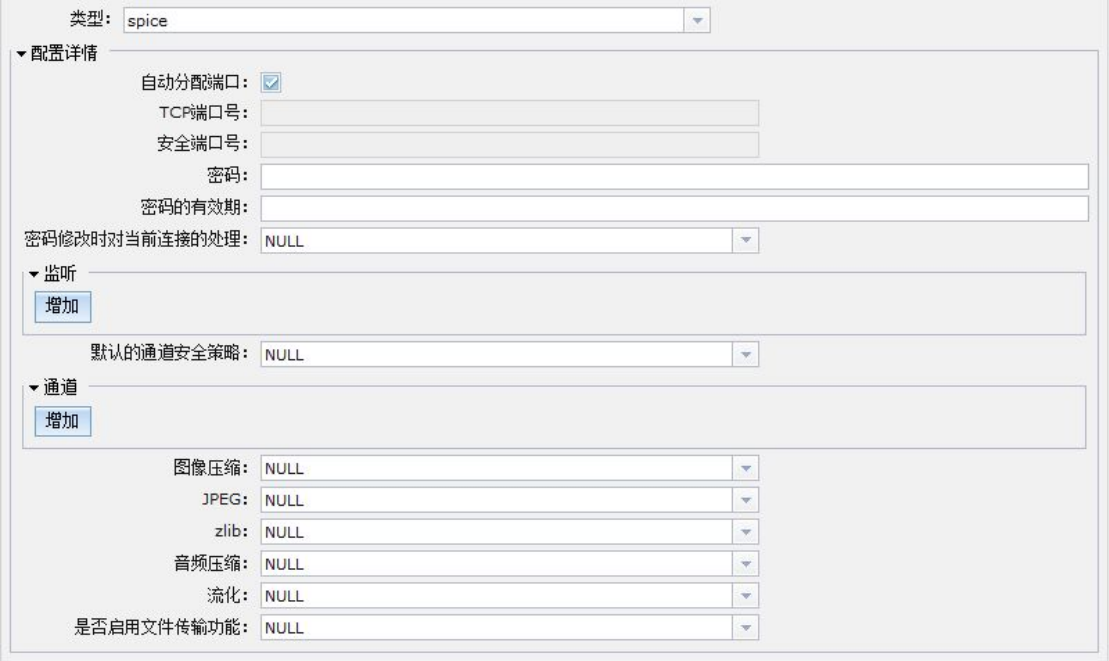


图 11- 9：图形界面 spice 配置

上面两图分别为“vnc”和“spice”的截图，各个配置项参数的填写请参考表“实例的图形界面配置”。

表格 11-3 实例的图形界面配置

编号	配置参数	选项说明	功能描述
1	远程桌面类型	目前只支持 vnc 和 spice 两种	远程桌面的服务方式。VNC 和 Spice 都是常用的方式。需要使用相应的远程桌面客户程序来访问服务。
2	是否自动分配端口	勾选框	勾选表示使用自动分配访问端口，否则表示使用指定的访问端口
3	TCP 端口号	TCP 通信端口号	非自动分配时指定的端口，使用从 5900 开始的端口，这里配置的是绝对值。
4	密码	访问虚拟机远程桌面服务的口令	为了安全，一般建议配置
5	密码的有效期	密码可以使用的最大期限	一般期限越短越安全
6	密码修改时保持当前连接	勾选框	勾选表示保持连接，否则表示断开连接

7	共享策略	目前支持以下的这些安全策略： allow-exclusive force-shared ignore	桌面显示共享策略，"allow-exclusive"允许客户端通过关闭其它连接来独占使用。要实现多个客户端同时连接则要求所有客户端都请求共享会话。这个值是默认值。"force-shared"禁止客户端独占使用，每个连接都必须为 vncviewer 指定-shared 选项。"ignore"无条件欢迎每个连接
8	监听：地址（IP）	IP 网络地址	提供远程桌面服务的监听地址。如果服务器有多个 IP 地址，而且希望只在某个 IP 地址上提供服务，则设置它。
9	安全端口号	TLS 通信端口号	Spice 特有的参数，使用 SSL 协议，数据加密通信的端口号
10	默认的通道安全策略	目前支持以下的选择：Any（默认值）、Insecure、Secure	Spice 特有，“Any”表示支持所有的连接；“Insecure”表示使用 TCP 通道，“Secure”表示使用 TLS 安全通道
11	通道：通道名称	通道名称	Spice 特有，Spice 支持多通道，英文和字母的字符串
12	通道：通道安全策略	目前支持以下的选择：Any（默认值）、Insecure、Secure	Spice 特有，“Any”表示支持所有的连接；“Insecure”表示使用 TCP 通道，“Secure”表示使用 TLS 安全通道
13	图像压缩	图像压缩接受 auto_glz，auto_lz，quic，glz，lz，off	图像的压缩模式
14	JPEG	图像的 JPEG 压缩接受 auto，never，always	“auto”表示自动压缩，“never”表示不压缩，“always”表示要压缩
15	zlib	zlib 配置广域网图像压缩接受 auto，never，always	“auto”表示自动压缩，“never”表示不压缩，“always”表示要压缩
16	音频压缩	接受为“开”，否则为“关”	音频流压缩，接受为“开”，否则为“关”
17	流化	流模式，为 filter，all 或 off 中的一个	“filter”表示部分，“all”表示全部接受，“off”表示全部拒绝
18	是否启用文件传输功能	开关量，选择“是”表示开，“否”表示关	文件传输功能，它默认是启用的，可以设置为“否”来禁用这个功能

第四步，配置虚拟机的网络。这里缺省没有网络，但是，虚拟机可以配置多个网络。

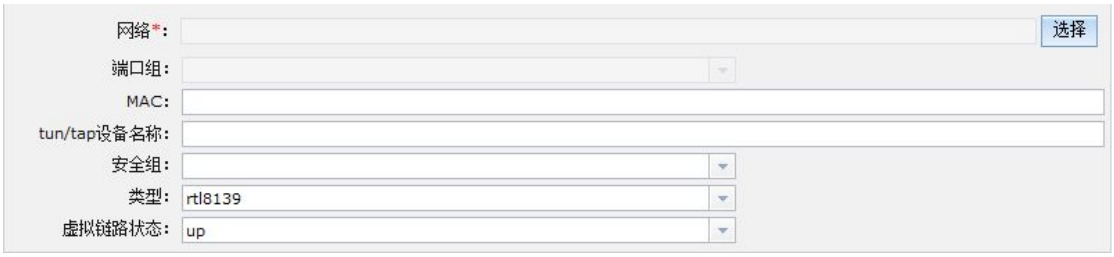


图 11- 10：网络配置

“网络配置”选项的参数的填写请参考表“实例的网络接口配置”。

表格 11-4 实例的网络接口配置

编号	配置参数	选项说明	功能描述
1	网络	选择需要使用的网络	这里的网络是由“网络管理”模块配置的虚拟网络。
2	端口组	选择网络的端口组	虚拟网络的端口组。并不是每个虚拟网络都配置有端口组，只有配置了的才可选。
3	MAC	MAC 地址	虚拟网卡的 MAC。如果不配置此选项，并且该虚拟机实例不是唯一的，则每次重启虚拟机后，MAC 地址将会重新自动生成，与前一次的地址可能不相同，具有随机性；如果配置了 MAC 地址，或者该虚拟机实例具有唯一性，那么每次重启虚拟机都将固定使用配置的地址，因此这时不能同时启动多个该虚拟机，否则将地址冲突。
4	tun/tap 设备名称	格式为 vnetN、tunN、tapN	Qemu 利用 tun/tap 设备作为虚拟机的网络设备。但不能保证一定是这个名称。
5	安全组	选择需要使用的安全组	安全组是网络过滤规则的集合，安全组由系统管理员制，在“网络管理”模块中配置
6	类型	模拟的网络接口类型。	设定要模拟的网络接口卡的模块
7	虚拟链路状态	开关量，选择“up”表示开，“down”表示关	如果指定 down，则接口就像网线断开了一样。默认是 up 状态

第五步，进行虚拟机的自定义配置。前面的配置可以满足绝大多数的情况，如果想要配置特殊的设备，可以在这里添加。添加优化信息时，必须在“domain”元素内；添加设备时，必须在“devices”元素内。

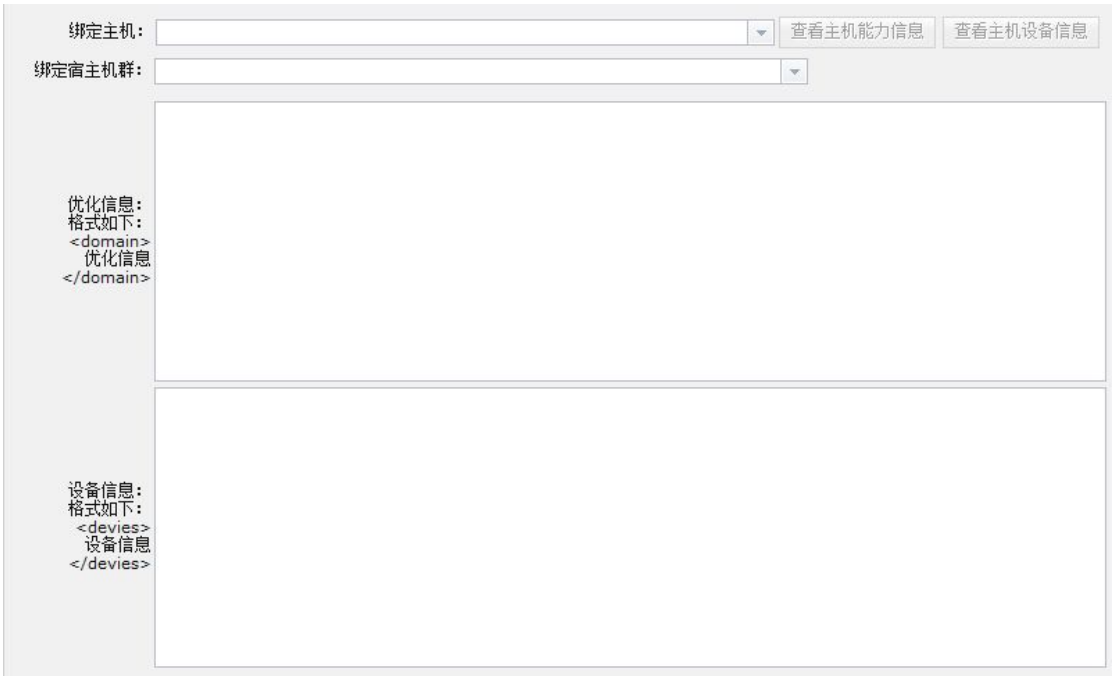


图 11- 11：自定义配置

上图是用户自定义实例部分的截图，各个配置选项的参数的填写请参考表“实例的用户自定义配置”。

表格 11-5 实例的用户自定义配置

编号	配置参数	选项说明	功能描述
1	绑定主机	选择需要绑定的主机	这里的主机是系统的服务器，它与“绑定宿主机群”是互斥的，二者只能选其一。
2	绑定宿主机群	选择需要绑定的主机	这里的宿主机群是在“宿主机群管理”模块处配置的，它与“绑定主机”是互斥的，二者只能选其一
3	优化信息	自定义的优化 XMl	格式必须要处于 domain 元素之内，例如： <domain> <features> <acpi default="on" toggle="yes"/> </features> </domain>
4	设备信息	自定义的设备 XMl	格式必须要处于 device 元素之内，例如： <devices> <video> <model type="vmvga" vram="10240" heads="1"/> <address type="pci" domain="0x0000" bus="0x00"

			slot="0x02" function="0x0"/> </video> </devices>

第六步，填写实例的中文名称，最后点击“保存”按钮提交。

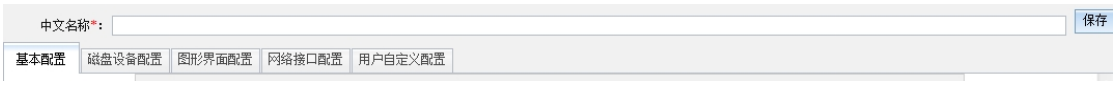


图 11- 12：保存虚拟机实例

在“虚拟机实例管理”界面上选中要修改的实例后，点击“修改”按钮，进入修改实例的界面，通过此界面可以查看实例的详细配置，也可以对其进行更改。该界面和新增虚拟机实例的界面完全是相同的，因此不再描述。

删除实例的方式是：先选中想要删除的实例，然后点击“删除”按钮，系统将弹出提示框，询问是否确定想要删除，如果确定，将会删除该实例，否则将取消。

禁用实例的方式是：先选中想要禁用的实例，然后点击“禁用”按钮，系统将弹出提示框，询问是否确定想要禁用，如果确定，将会禁用该实例，否则将取消。被禁用的虚拟机实例被标志为不再使用。如果该实例还有虚拟机在运行，系统将强制关闭它们。而且，选择实例时，“管理虚拟机”的按钮不再能使用，即无法再管理该实例的虚拟机。

激活实例的方式是：先选中想要激活的实例，然后点击“激活”按钮，系统将弹出提示框，询问是否确定想要激活，如果确定，将会激活该实例，否则将取消。只有处于激活状态的虚拟机实例才能被使用。

经测试，数字有机体虚拟机系统支持大量的主流的操作系统，如 windowsXP、windows-server-2008、windows7、windows8、windows10、neokylin、startos、debian、ubantu、suse、redhat、fedora、centos、freebsd、pcbsd 和 openbsd 等。

以下几个表的内容是一个完整的虚拟机实例的配置示例。该虚拟机定义为 64 位计算机，定义了虚拟的硬盘和虚拟的光盘，虚拟的光盘中指定了“windowsXP”操作系统的安装镜像。

表格 11-6 基本配置示例

编号	输入项	值	备注
1	机器类型	x86_64	表示模拟 64 位计算机
2	虚拟机监控程序	kvm	
3	虚拟机类型	pc	
4	虚拟机名	test-vm	
5	启动菜单-是否启用	是	启动时可以选择启动设备
6	启动菜单- 启动超时时间	10000	长时间方可选择启动设备

7	虚拟 CPU 数量	1	
8	内存分配-内存分配容量	1	
9	内存分配-单位	GB	

表格 11-7 磁盘设备的硬盘示例

编号	输入项	值	备注
1	资源类型	volume	
2	缓存模式	writeback	
3	读时复制	关	如果读时复制配置为“开”，制作包含该磁盘的检查点时将失败
4	存储池	Peng-pool	用户可以根据需要修改为已有的存储池
5	存储卷	Peng-vol12	用户可以根据需要选择使用的存储池中的存储卷
6	启动故障策略	Optional	
7	总线	Sata	Window 系统建议使用 sata 或者 IDE。
8	只读	未选择	通常只对光盘设置为只读。

表格 11-8 磁盘设备的光盘示例

编号	输入项	值	备注
1	资源类型	file	
2	设备类型	cdrom	这里定义为安装光盘
3	镜像文件	/dpfs/winxp-ghost.iso	这里使用的是已经上传的 windowsXP 光盘的镜像文件
4	启动故障策略	mandatory	
5	重标安全标签	开	
6	总线	ide	
7	托盘状态	收回	
8	只读	选择	光盘一般都是只读的

表格 11-9 图形界面配置的示例

编号	输入项	值	备注
1	类型	vnc	
2	是否自动分配端口	选择	为防止端口冲突，这里最好采用自动分配端口
3	密码	空值	
4	密码的有效期	空值	
5	密码修改时是否保持当前连接	未选择	

6	共享策略	ignore	
7	监听	0.0.0.0	监听所有的接口地址

表格 11-10 网络接口配置的示例

编号	输入项	值	备注
1	网络	Linux-br-network	配置的虚拟网络需要先配置。其网络类型为“Linux 桥接”，“转发网络接口”定义为“eth1”。
2	端口组	空值	
3	MAC	空值	
4	tun/tap 设备名称	空值	
5	安全组	空值	
6	类型	rtl8139	
7	虚拟链路状态	up	

表格 11-11 用户自定义配置的示例

编号	输入项	值	备注
1	绑定主机	空值	
2	绑定宿主机群	空值	
3	优化信息	<domain> <features> <acpi/> </features> </domain>	此值是自动生成的
4	设备信息	<devices> <serial type="pty"> <source path="/dev/pts/1"/> <target port="0"/> </serial> <video> <model type="vmvga" vram="10240" heads="1"/> <address type="pci" domain="0x0000" bus="0x00" slot="0x02" function="0x0"/> </video> </devices>	此值是自动生成的

11.4 虚拟机管理

在虚拟机实例管理界面上，虚拟机实例的信息是以表格行来展示的，表格行描述了虚拟机实例的中英文名称、绑定的宿主机、虚拟的 CPU 数量、虚拟的内存容量、拥有的虚拟机数量、管理状态和创建日期。点击“修改”按钮可以查看每个虚拟机实例的更详尽信息。当虚拟机实例数量过多时，表格将以分页的形式显示，支持按中英文名称搜索。

在虚拟机实例管理界面上，先选中想要管理的“虚拟机实例”，然后点击“管理虚拟机”按钮，就能够进入“虚拟机管理”界面。虚拟机管理界面如图 11-13 所示。

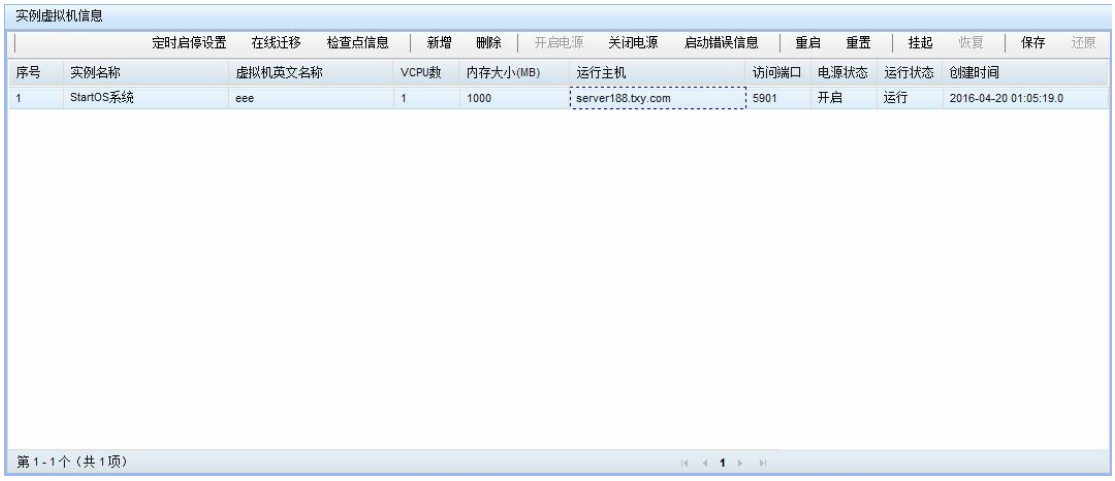


图 11- 13：虚拟机管理界面

管理界面主要包括以下这些操作：

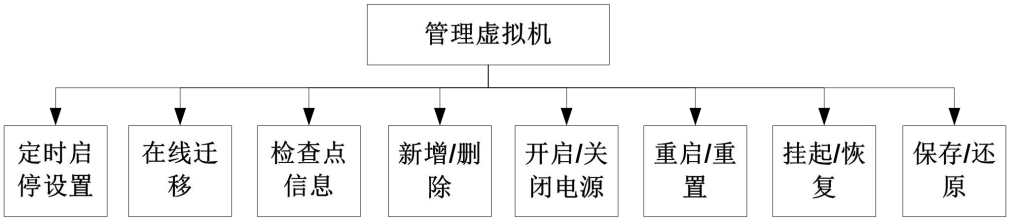


图 11- 14：管理虚拟机

11.4.1 定时启停

本小节将介绍虚拟机的定时器。虚拟机的定时器为用户提供了一种简易的控制虚拟机启停的方式。它允许用户设置三种虚拟机的启停方案：

1) 按照一定周期的循环定时操作（包括启动、停止、运行时间段、停止时间段）；
2) 单次的预约运行时间段；
3) 固定的间隔操作（包括运行时间间隔和停止时间间隔）。

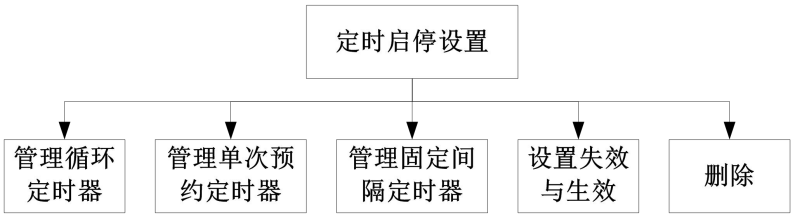


图 11- 15：虚拟机的定时启停

基于以上三种方案，几乎能满足大多数的运行计划。为了保证各个启停计划的顺利进行，虚拟机的启停计划时间是不能冲突的，即时间不能重复。首次使用定时器，定时计划是没有的，如图 11-16 所示为空的“虚拟机定时启停”界面：

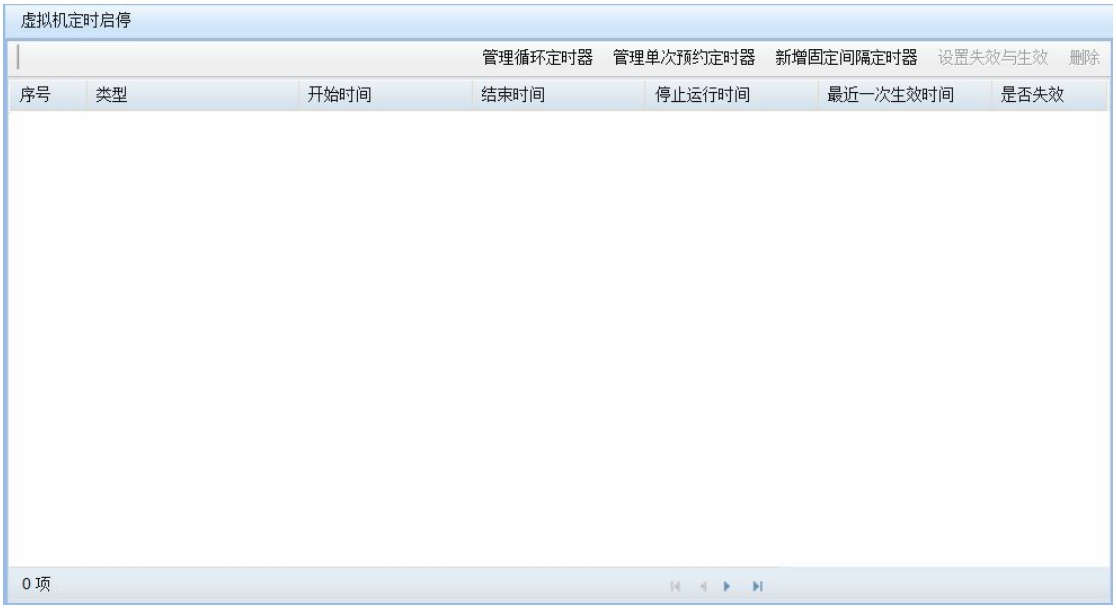


图 11- 16：无设置时的虚拟机定时启停管理界面

11.4.1.1 管理循环定时器

图 11-17 是（周期）循环定时器的编辑界面。该编辑界面能够“新增”定时计划和“修改”已有的定时计划。增加定时计划的方法是：按住 Ctrl 键，并在网格上拖动鼠标。但是需要注意的是：新增加计划时需要先在编辑界面的左边栏设定“定时器类型”和“周期”，新增后可以编辑左边栏的“时间计划名称”、“开始时间”和“结束时间”，当然也可以通过拖动时间计划来确定开始结束时间。最后，还可以点击“更新内存数据和显示”按钮来刷新编辑界面。

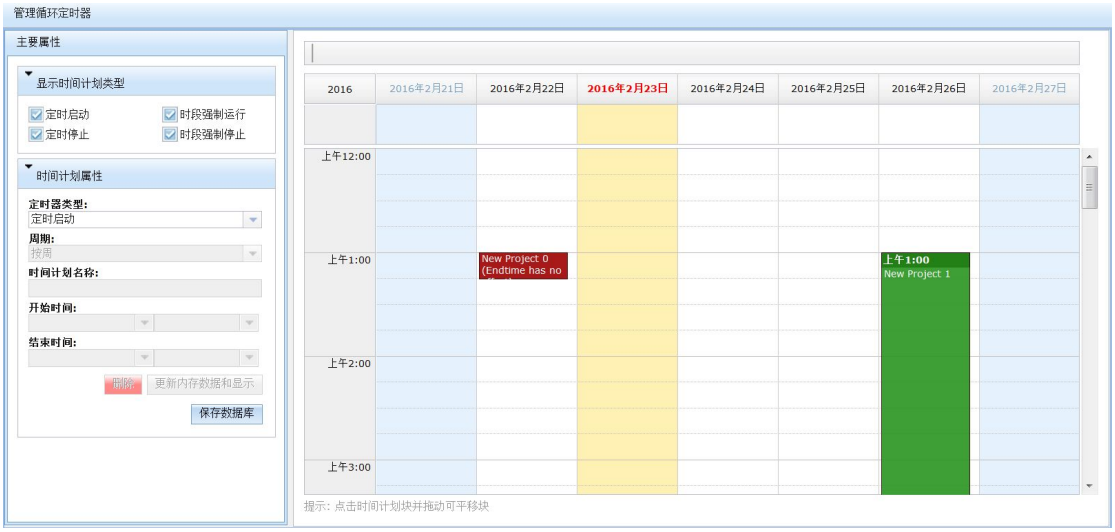


图 11- 17：管理循环定时器

不同周期的定时计划存在冲突的可能性极大，也不好控制，因此本系统在使用了某一种周期后就不能再使用其他周期的定时计划。为了能够显著地分辨出“定时器类型”的四种类型，本系统采用绿色表示“时段强制运行”，红色表示“定时启动”，蓝色表示“定时停止”，黄色表示“时段强制停止”。图 11-18 所示为四种不同类型的定时计划，如果不想显示某一类型的定时计划，可以在编辑界面的左上角将该定时计划复选框反选。定时计划的时间可以通过鼠标拖动来改变，也可以直接编辑“时间计划属性”，点击“保存数据库”按钮后起效。

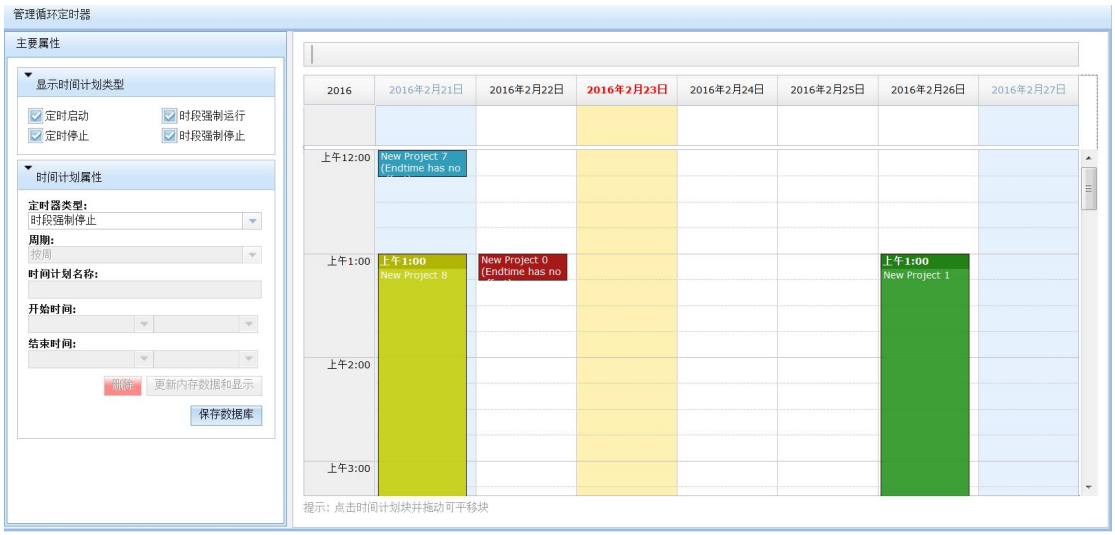


图 11- 18：循环定时计划示例

表格 11-12 时间计划属性

编号	输入项	输入内容
1	定时器类型	“定时器类型”包括“定时启动”、“定时停止”、“时段强制运行”和“时段强制停止”四种类型
2	周期	“周期”包括“按天”、“按周”和“按月”三种类型
3	时间计划名称	字符串，用于标识该定时计划的名称
4	开始时间	定时计划的开始时间

5	结束时间	定时计划的结束时间

添加了定时计划后，在“虚拟机定时启停”界面上就能查询到了，如图 11-19 所示。再次点击“管理循环定时器”按钮，可以查看已有的定时计划，也可以对其进行修改。

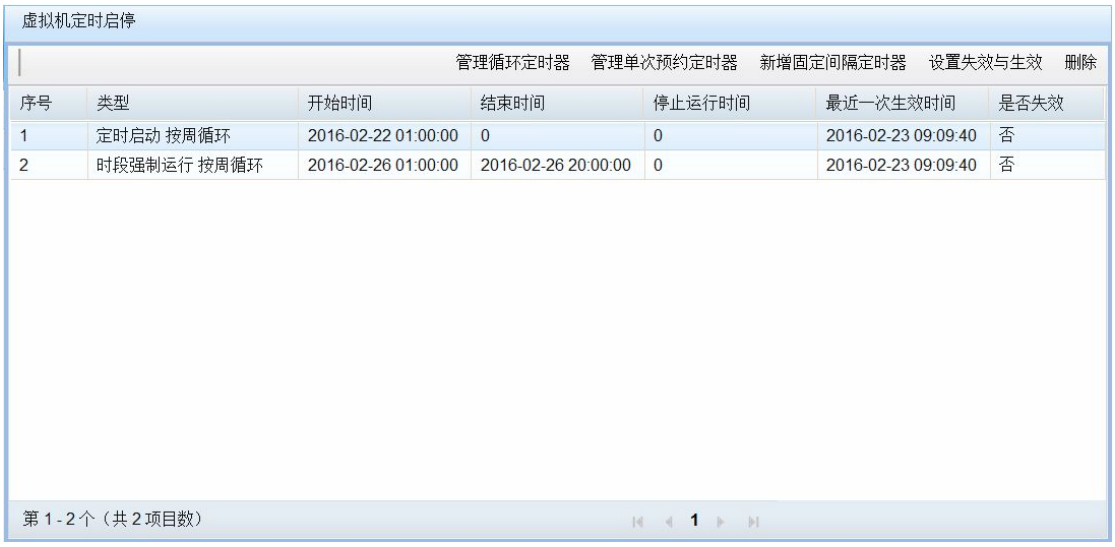


图 11- 19：循环定时计划列表

11.4.1.2 管理单次预约定时器

单次预约是不会周期循环的，每个计划只会运行一次。以下为进入单次预约的编辑界面，该界面是以日历的形式展示的，用户可以先按住“Ctrl”键，再在日历上点击来新建定时器的预约计划，这里的预约计划只有一种类型，就是一个运行的时间段（包括一个开始时间和一个结束时间），开始和结束时间可以通过拖动时间计划来改变，也可以直接编辑“时间计划属性”的“开始时间”和“结束时间”参数。

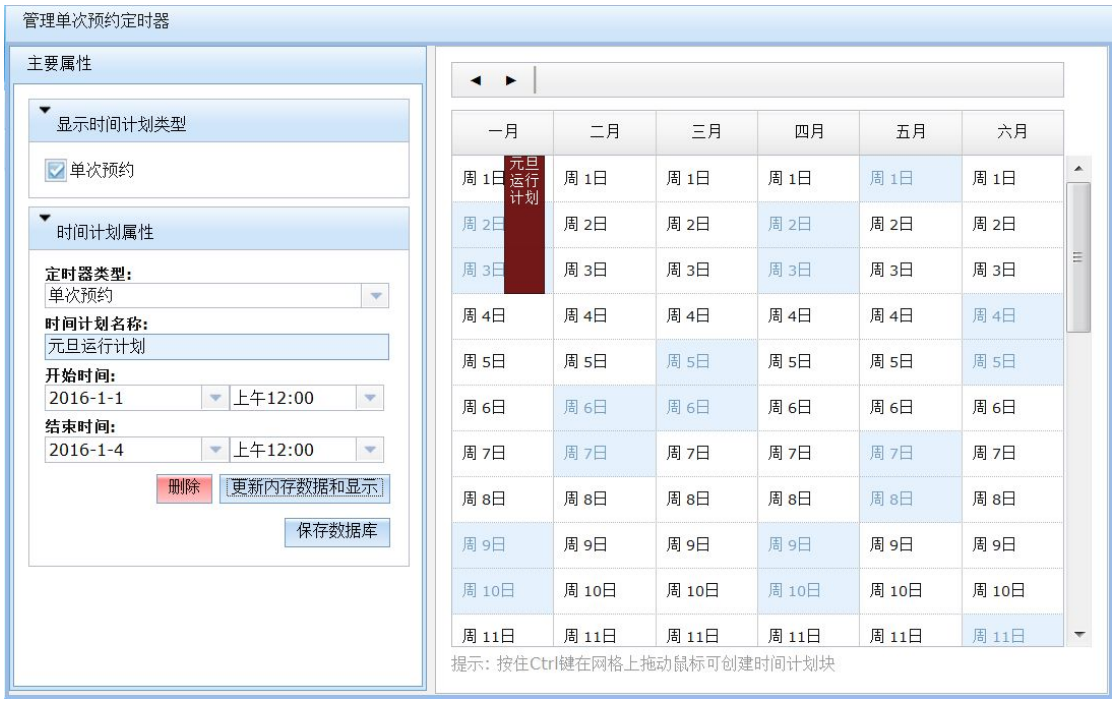
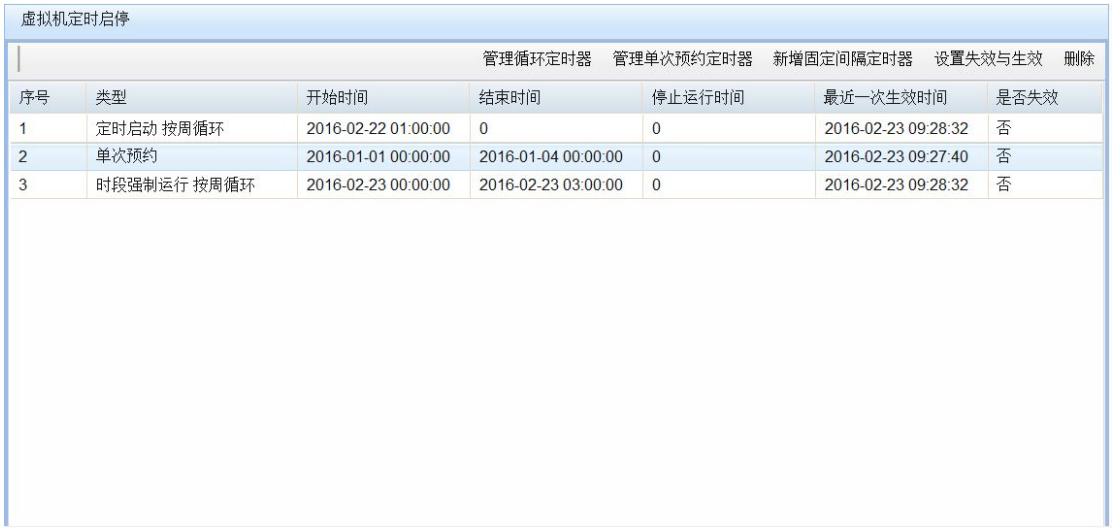


图 11- 20：管理单次预约定时器

单次预约定时计划添加成功后，也可在“虚拟机定时启停”界面上查询到，如图 11-21 所示的选中行，属于增加的单次预约定时计划。



序号	类型	开始时间	结束时间	停止运行时间	最近一次生效时间	是否失效
1	定时启动 按周循环	2016-02-22 01:00:00	0	0	2016-02-23 09:28:32	否
2	单次预约	2016-01-01 00:00:00	2016-01-04 00:00:00	0	2016-02-23 09:27:40	否
3	时段强制运行 按周循环	2016-02-23 00:00:00	2016-02-23 03:00:00	0	2016-02-23 09:28:32	否

图 11- 21:添加单次预约后的定时器列表

11.4.1.3 固定间隔定时器

固定间隔定时器是这样的一种定时器：设置一个起始时间，然后再设置一个运行的时间段和一个停止运行的时间段。这样，虚拟机就会从起始时间开始总是先运行一段时间，然后停止运行一段时间，之后又重复开始运行一段时间，周而复始。如下图所示，设置固定间隔定时计划的各个选项都是必须要填写的，运行时间段和停止时间段以“分钟”为单位。

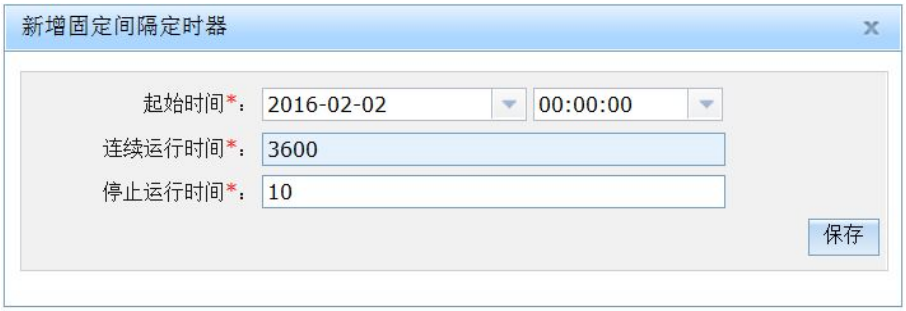


图 11- 22：固定间隔定时器

以下为试着设置一个固定定时计划的操作，最终操作失败（图 11-23），原因是本系统目前尚不支持“循环定时器”和“固定间隔定时器”同时存在的情况。想要设置“固定定时器”，需要先删除“循环定时器”中的计划。如果新建成功，也可在“虚拟机定时启停”界面上查询到。

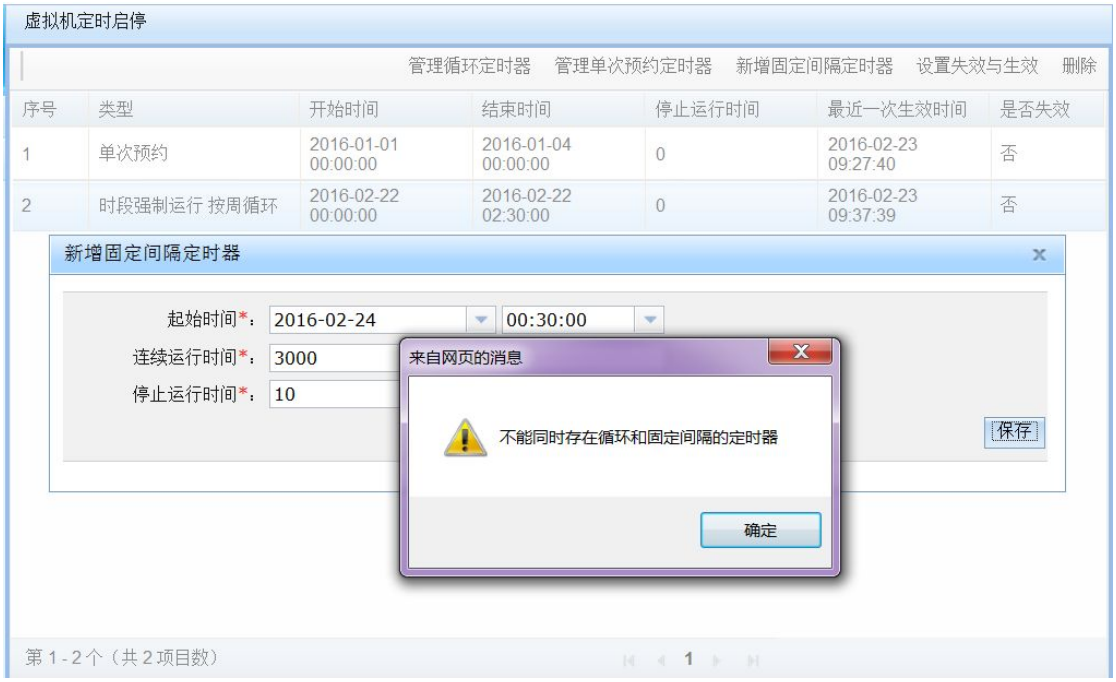


图 11- 23：设置固定间隔定时器失败的提示

11.4.1.4 定时器有效性管理与删除

定时计划新建成功后都能够在“虚拟机定时启停”界面上找到。当暂时不再需要时，可以将其设定为无效，需要使用的时候再将其设定为有效。当永远不再需要时，还可以选择删除该定时计划。图 11-24 为点击“设置失效与生效”按钮后对话框，选择编辑是否失效下拉框，最后点击确定按钮就可更改定时计划的有效性。

图 11- 24：设置定时器状态的选择界面

如果不再需要某定时计划，可以点击“删除”按钮。图 11-25 为是否删除的询问提示，选择“确定”后将删除，否则取消。

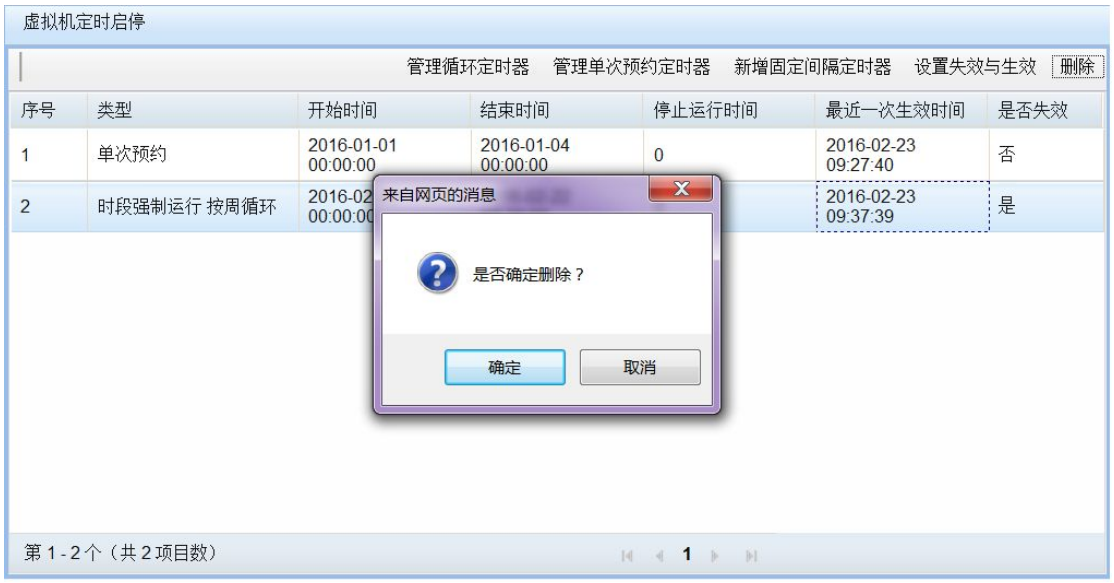


图 11- 25：删除定时器的提示

11.4.2 在线迁移

虚拟机的迁移，就是把一个虚拟机从一台服务器转移到另一台服务器。通过“关闭电源”和“开启电源”两个按钮的配合能够实现虚拟机的离线迁移，即先关闭服务器 A 上的虚拟机，再在服务器 B 上启动虚拟机。这样的迁移可以解决一部分问题，但是如果用户有更加严格的需求，比如想要在不关闭虚拟机的前提下实现虚拟机的迁移。我们把这种不关闭虚拟机的前提下实现虚拟机的迁移称作在线迁移，也被称作“热迁移”。

在线迁移存在一定的限制。如果用户配置了虚拟机可以运行的宿主机群，则虚拟机可以在宿主机群内迁移。如果虚拟机绑定了宿主机，则虚拟机不能迁移。如果用户既没有绑定宿主机，也没有配置宿主机群，则虚拟机可以任意迁移。

任何时候都可以迁移虚拟机。当虚拟机处于未运行、挂起或者保存状态时，迁移虚拟机仅仅是改变虚拟机下次运行的位置，并不会产生什么影响。但是，对正在运行的虚拟机来说，迁移后打开的网络连接将丢失，其余都可继续保持。

用户可以根据需要手动地迁移虚拟机。这时用户通过管理页面发出指令，由当前控制虚拟机的宿主机完成迁移工作（当然需要迁移的目的主机的配合）。有一个问题是：系统能否自动的迁移虚拟机。显然，自动迁移的主要目的是负载均衡。当新增加服务器，或者因某些虚拟机停止运行而导致负载不均衡时，有必要迁移虚拟机以均衡负载。目前系统支持手动的在线迁移，不支持基于负载的自动在线迁移，但是由于“虚拟机集群”的需要，系统使用自动离线迁移虚拟机来实现负载均衡。

数字有机体虚拟机系统具备在线迁移的功能，但是在不同 CPU 结构的服务器迁移时可能会失败。一般都是由于目标服务器和源服务器的配置存在差异造成的，也有可能因资源被占用而失败。此外，虚拟机的虚拟磁盘使用的卷和镜像都必须存放在共享文件系统中，只有源服务器和目标服务器都能访问才能进行在线迁移。如果资源不是共享的，但是又想要实现在线迁移，可以手动将源服务器上的本地资源拷贝到目标服务器上相同位置，还可以使用 NFS 将源服务器的资源共享到目标服务器的相同位置，然后再发起“在线迁移”。

图 11-26 显示为在线迁移前的服务器运行状态，状态显示虚拟机运行在服务器

“server185.txy.com”上，VNC 监听端口为 5900。

图 11- 26：迁移前虚拟机状态

使用服务器“server185.txy.com”的 IP 地址和端口 5901 连接 VNC 服务器，显示为一个 windows7 的系统界面，如图 11-27 所示。

图 11- 27：迁移前的虚拟机

点击“在线迁移”按钮，系统弹出“迁移宿主机选择”对话框，在此对话框中选择迁移的目标服务器（比如这里选择 server188.txy.com 服务器），最后点击“确定”按钮，如图 11-28 所示，之后系统开始迁移虚拟机。

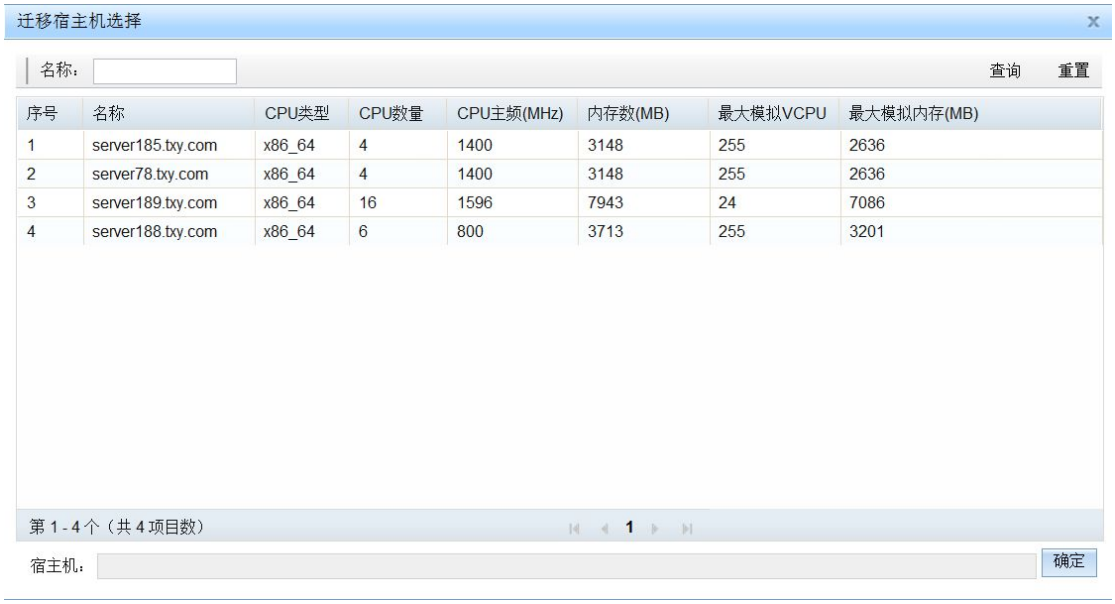


图 11- 28：在线迁移虚拟机

迁移成功后，状态显示虚拟机运行在服务器“server188.txy.com”上，VNC 监听端口为 5901，如图 11-29 所示。

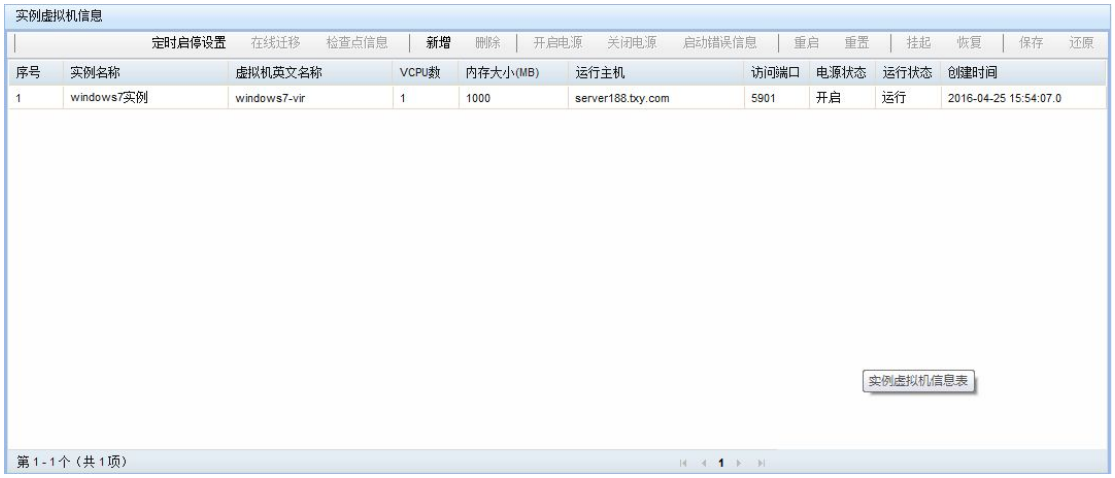


图 11- 29：迁移后的虚拟机状态

使用服务器“server188.txy.com”的 IP 地址和端口 5901 连接 VNC 服务器，显示和迁移前的系统界面完全一致，说明虚拟机迁移成功。

11.4.3 检查点信息

检查点对虚拟机来说具有很重要的作用。利用检查点的特性，可以防止虚拟机的数据丢失，可以回退到以往的虚拟机运行状态。检查点是虚拟机运行时的一个快照，此快照能够同时记录制作检查点之时的内存和磁盘状态。制作的检查点是保存在用户指定目录中的，各个检查点之间的关系呈树形结构。

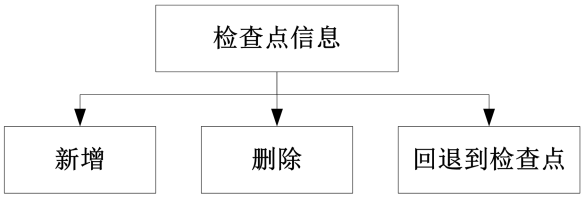


图 11- 30：检查点管理

检查点的管理主要包括新增检查点、删除检查点和修改检查点。删除检查点时，需要先选择删除对象，然后再删除。删除的对象可以是树的叶子，也可以是非叶子。当选择非叶子时，将自动删除其下的所有节点。图 11-31 为检查点管理界面，展示的树形结构如下：

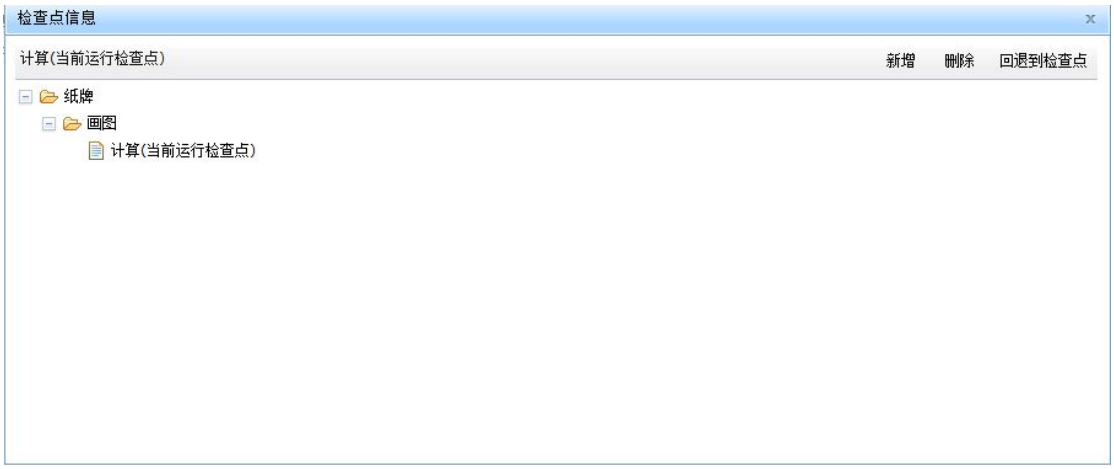


图 11- 31：检查点的管理界面

点击“新建”按钮，弹出新建检查点的用户界面。需要填写检查点的“中文名称”、“英文名称”、“保存路径”、“保存检查点的文件夹名”，这些选项都是必须要填写的。如果该虚拟机存在磁盘，作为可选项，还可以选择想要制作快照的磁盘。如果系统中没有磁盘，则“选择磁盘”这个可选项为空，如图 11-32 所示。

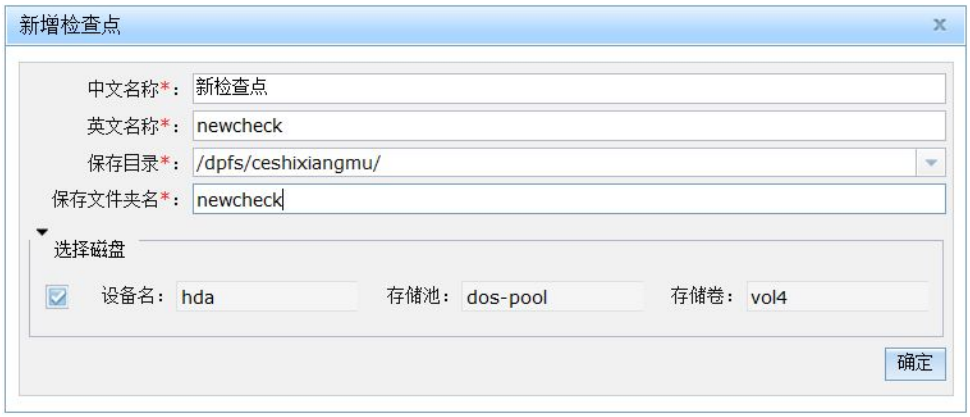


图 11- 32：新建检查点

填写完毕信息，点击“确定”按钮后，开始新建检查点，如果新建成功，新的检查点是检查点树型结构的新的叶子，如图 11-33 所示，被选中的节点为新增的检查点。

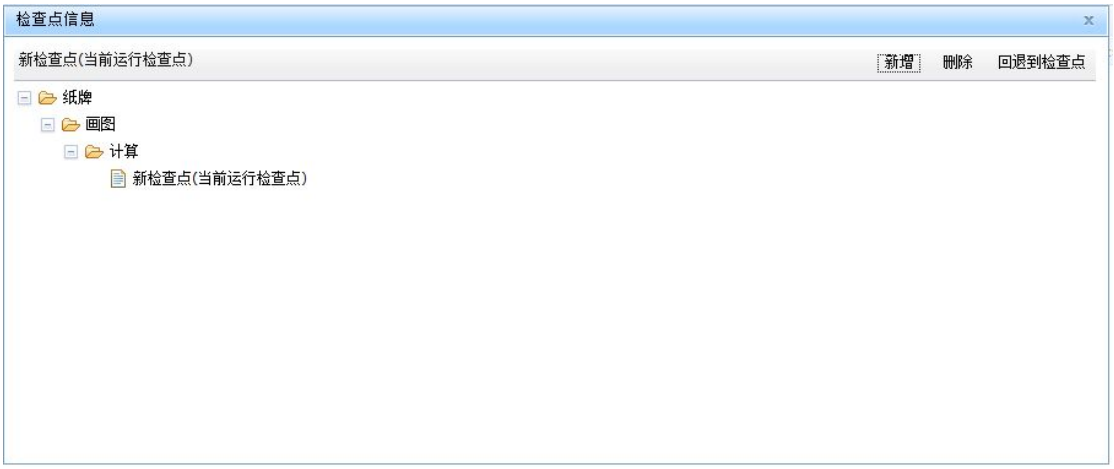


图 11- 33：新建检查点成功

如图 11-33 所示，已经有检查点“纸牌”、“画图”、“计算”、“新检查点”，由于最后一次制作的检查点是“新检查点”，因此当前的检查点是“新检查点”。虚拟机当前显示的桌面如图 11-34 所示。

图 11- 34：回退前的虚拟机画面

现在，我们想要回退到“画图”，操作的方法是：首先选择“画图”，然后点击“回退到检查点”按钮。如所示，系统会询问用户“是否确定回退检查点”。如果确定，就将执行回退任务，否则取消。

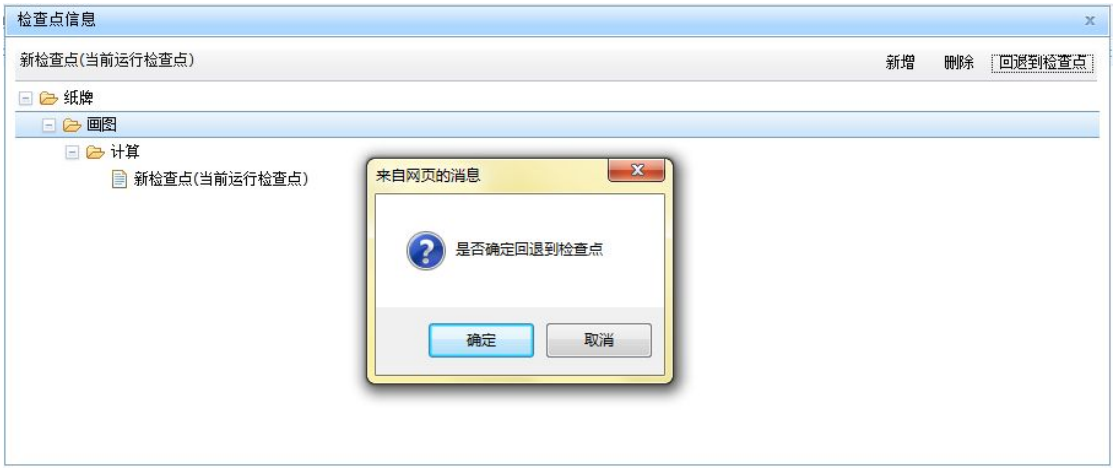


图 11- 35：回退检查点提示

点击询问消息的“确定”按钮，开始回退操作，回退可能需要花费一些时间，回退成功后，系统提示“操作成功”。这时，通过远程桌面可以看到虚拟机又回到了检查点的运行位置，下图是回退后的虚拟机画面。

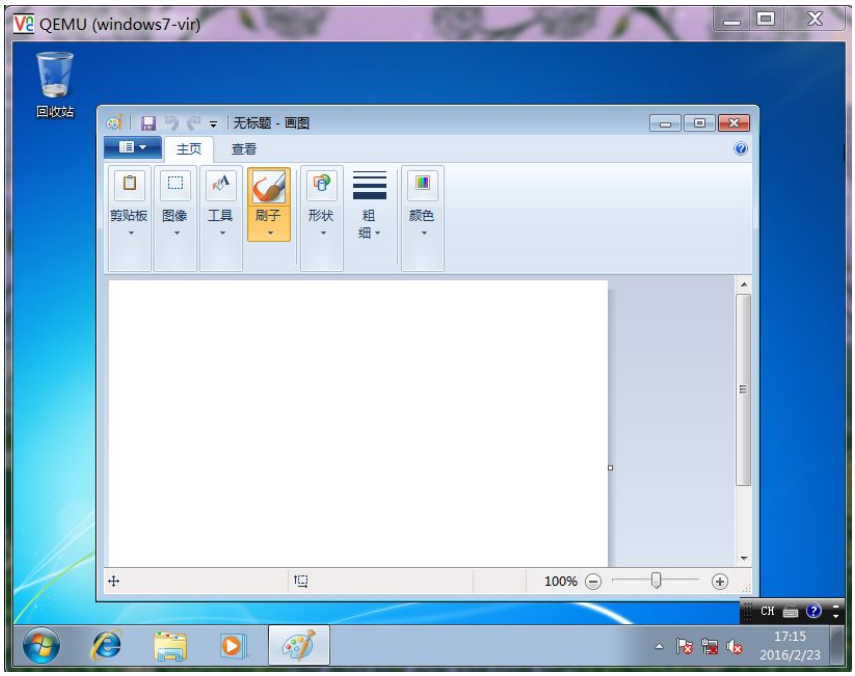


图 11- 36：虚拟机回退后的画面

所有的检查点是以树形结构显示的。因此删除检查点时，可以删除树的叶子节点，也可以删除非叶子节点。删除前先要选择想要删除的节点。如果选择的是叶子节点，表示只删除当前选择的项；如果选择的是非叶子节点，则表示删除包括选择项在内，以及其下的所有项。删除的方法是：选择想要删除的检查点节点（如图 11-37 选中了“计算”这个检查点），然后点击删除，点击询问提示框的“确定”按钮。

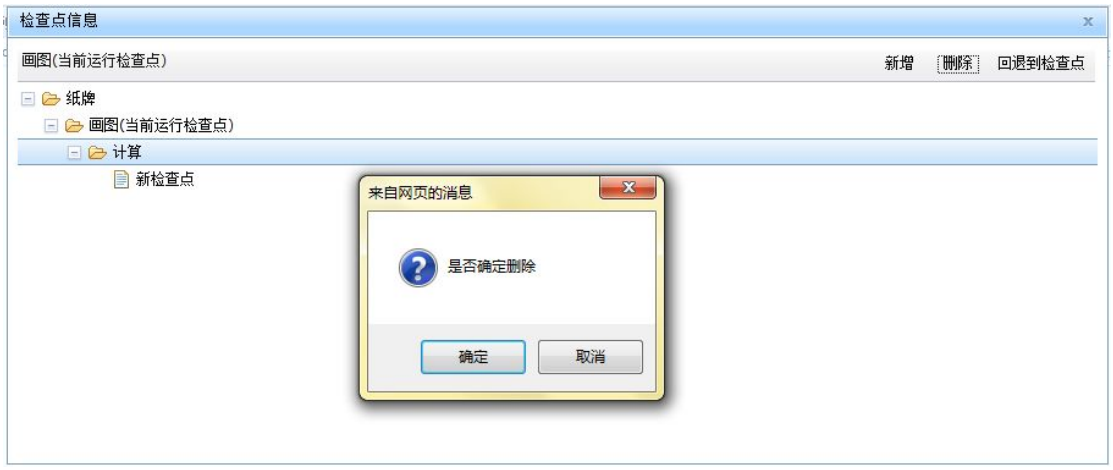


图 11- 37：删除检查点的界面

删除检查点需要较长的时间，成功删除后将给出操作成功的提示信息。再次查看检查点将看到已经删除的检查点不再存在了。

11.4.4 新增和删除虚拟机

以下将介绍虚拟机的增加与删除。虚拟机的增加方式有两种，一种方式是显示地增加，一种方式是虚拟机集群增加的。虚拟机集群中，可以设置使用虚拟机的数量，也可以根据负载自动增加虚拟机，详细的情况请参阅虚拟机集群的章节。这里将着重介绍显示的增加方式。

如果从一个实例新建多个虚拟机，并且多个虚拟机可能会同时运行的时候，应该保证虚拟机实例的存储设备（如磁盘和 cd 设备）应该是“只读”的。多个虚拟机同时使用“可写”的存储设备可能会出现错误。因此，虚拟机实例表中显示为“唯一虚拟机”的实例只能新建一个虚拟机。

以下截图显示名称为“peng_DOS 集群系统”的虚拟机实例有一个名称叫“server100”的虚拟机。接下来我们将新增两个虚拟机。

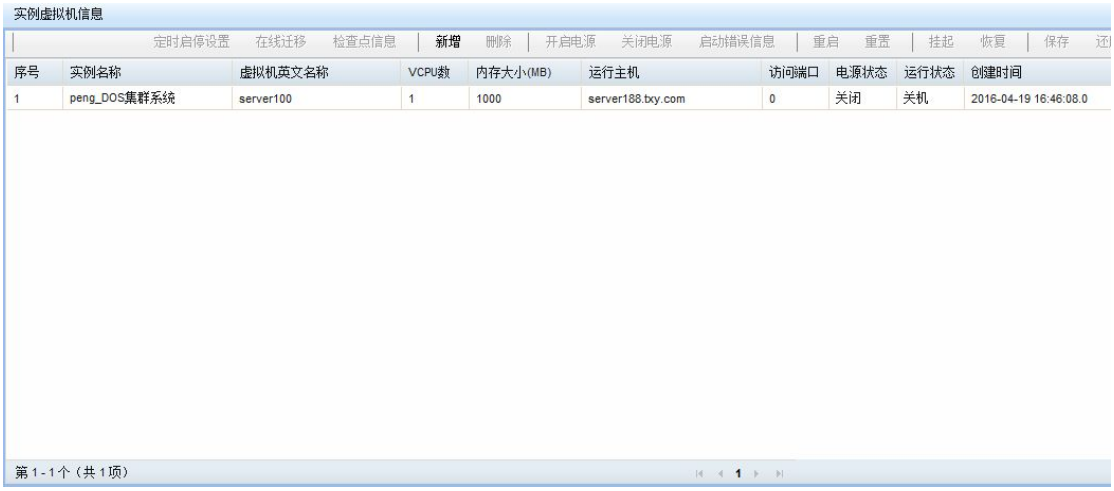


图 11- 38：新增虚拟机前实例的虚拟机情况

点击“新增”按钮，系统将弹出“新增虚拟机”的对话框，在对话框中输入“英文名”

（如这里输入“vvm5”），此项为必选项，并选择是否“立即启动”。最后点击“确定”按钮。

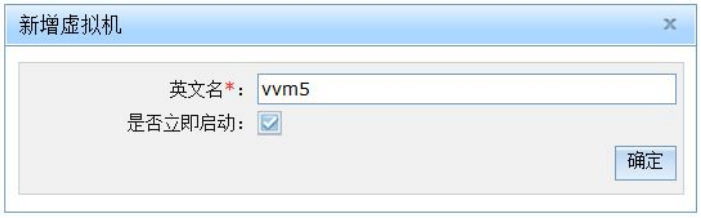


图 11- 39：新增虚拟机界面

再点击“新增”按钮，系统将弹出“新增虚拟机”的对话框，在对话框中输入“英文名”（如这里输入“vvm6”），此项为必选项，并选择是否“立即启动”。最后点击“确定”按钮。

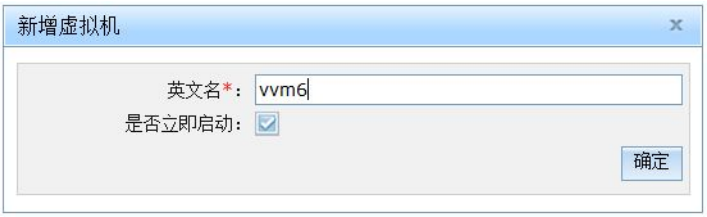


图 11- 40：新增另一个虚拟机

经过前面两次的新增虚拟机操作后，现在一共有三个虚拟机，如图 11-41 所示。

图 11- 41：新增虚拟机成功后实例的虚拟机情况

删除虚拟机的方法是：先选中想要删除的虚拟机（如“vvm5”），然后点击“删除”按钮，如图 11-42 所示，在弹出的询问提示框中选择“确定”按钮就将删除选中的虚拟机。删除了的虚拟机将不再实例的虚拟机列表中显示。

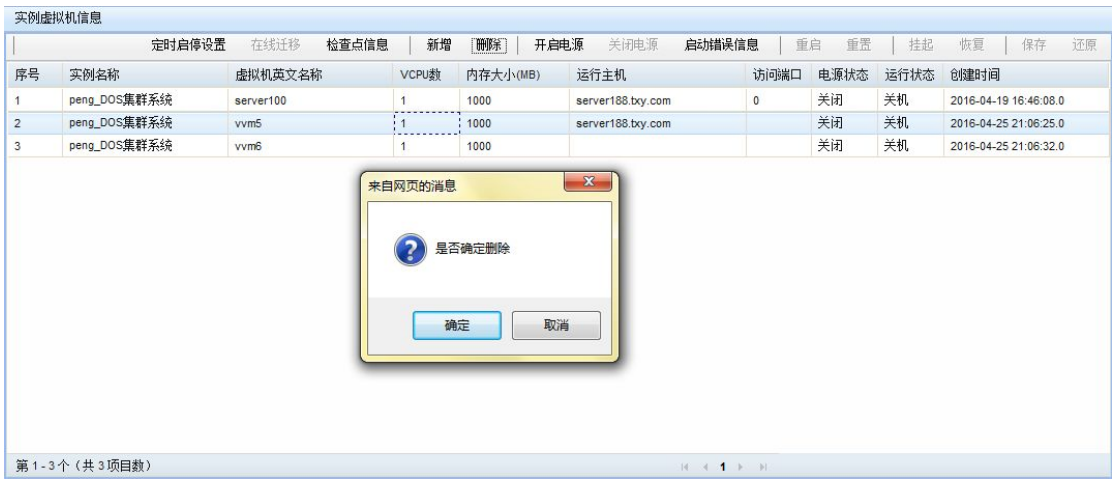


图 11- 42：删除虚拟机

11.4.5 开启和关闭虚拟机

对开启和关闭虚拟机，既支持按照计划的自动的定时开启和关闭机制，也支持手动的开启和关闭操作。本手册在“定时启停”章节已经阐述了如何设置定时启动和关闭，以下将介绍虚拟机的手动开启和关闭。虚拟机的开启按钮是“开启电源”，关闭按钮是“关闭电源”。点击开启虚拟机按钮后，系统会尝试开启虚拟机，虚拟机的运行状态显示为“运行”时，表示已经启动，显示为“关机”时，表示尚未启动，或者启动未成功。同样，点击关闭虚拟机按钮后，系统会尝试关闭运行中的虚拟机，关闭是否成功可以查看运行状态。处于电源“开启”状态的虚拟机总是优先排列在“实例虚拟机信息表”的开头位置。

开启时，如果一个实例的多个虚拟机在同时运行的时候，应该保证虚拟机实例的存储设备（如磁盘和 cd 设备）应该是“只读”的。多个虚拟机同时使用“可写”的存储设备可能会出现错误。

开启虚拟机的时候可以选择运行的主机，也可以不选择，不选择表示系统自动调度。能够选择的宿主机根据为实例设定的范围来确定。如果实例绑定了运行主机或者宿主机群，则只显示相应的宿主机。如下图所示为点击“开启电源”后弹出的选运行主机对话框。选择一个主机后确定。该虚拟机就将在选择的主机上运行。如果虚拟机启动的过程中发生错误，则通过点击“启动错误信息”来获取出错消息。

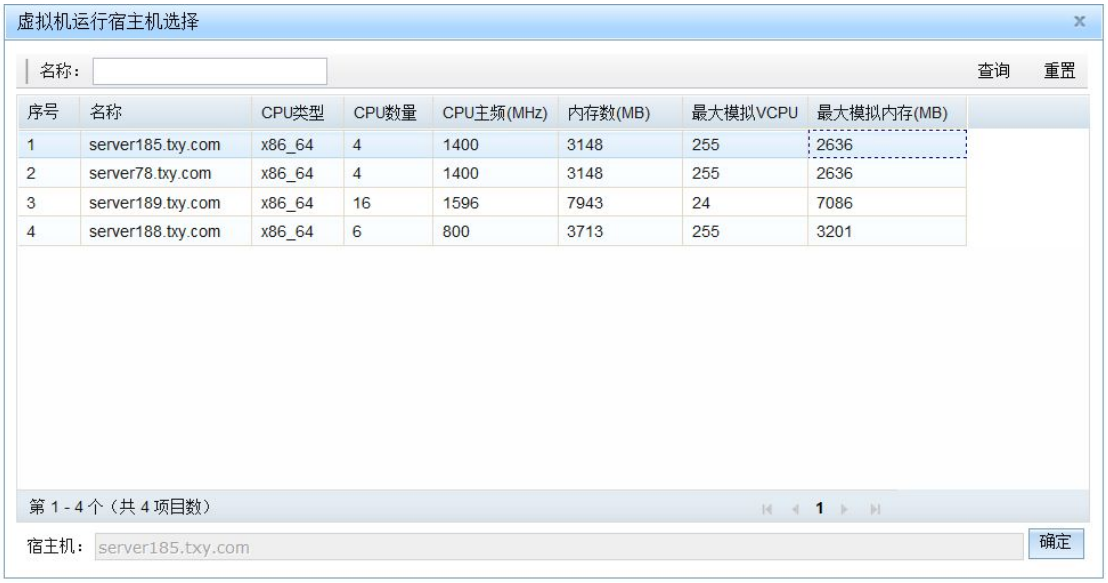


图 11- 43：选择虚拟机运行属主机

该操作通常会很快报成功，但虚拟机还需要等待具体的宿主机将其启动起来。可以通过刷新实例的虚拟机列表来查看要启动的虚拟机是否真正启动成功。虚拟机真正启动成功后，在虚拟机列表中，其“运行状态”将由原来的“关机”变成“运行”，并显示当前运行的主机，VNC 端口也将有具体的值，如 5900。

为了确认该虚拟机在正常的运行，可通过 VNC 客户端远程访问虚拟机。在 VNC 终端中输入当前运行宿主机的 IP 地址和 VNC 端口进行连接。图 11-44 为使用 VNC 连接的输入截图。

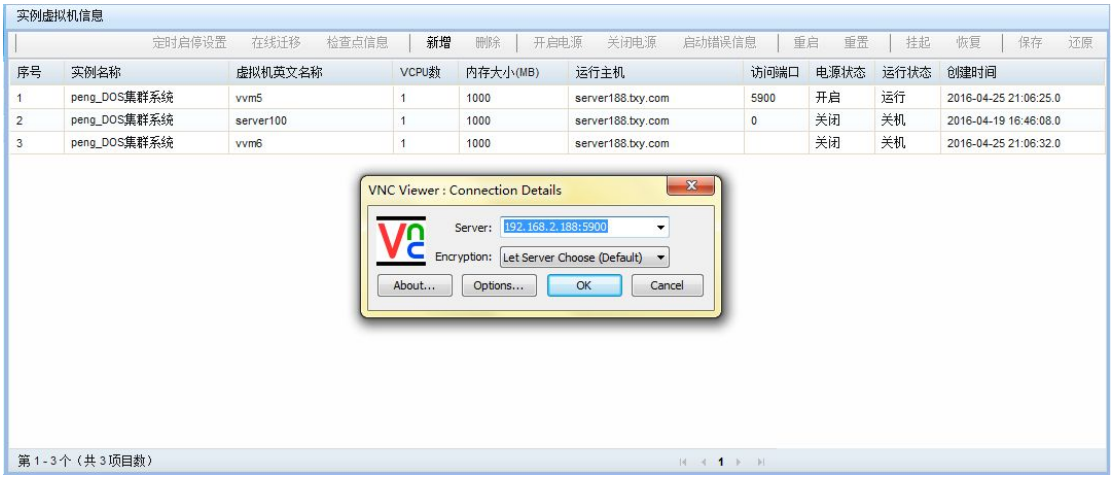


图 11- 44：使用 VNC 访问正在运行的虚拟机

连接 VNC 成功，并显示出了“数字有机体系统”的登录界面，说明系统确实处于运行之中。

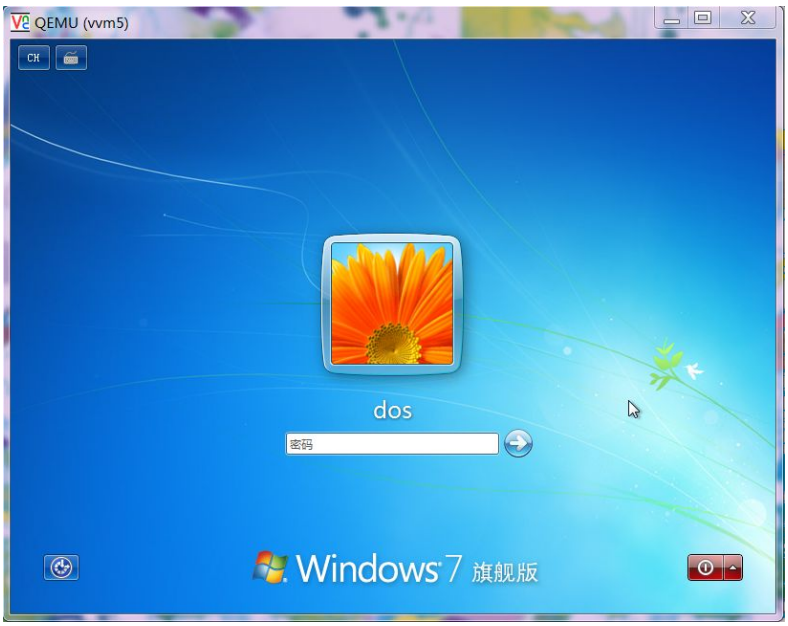


图 11- 45：用 VNC 远程访问到的虚拟机画面

前面介绍了如何开启虚拟机，接下来将介绍如何关闭虚拟机。在数字有机体虚拟机系统中，直接到服务器上去关闭虚拟机是无法关闭它的，必须通过管理网站才能完整地关闭虚拟机。关闭的方式是：先选择想要关闭的虚拟机，然后再点击“关闭电源”按钮。点击“确定”按钮后将关闭虚拟机。如果虚拟机正确的配置了电源管理，则虚拟机的关闭将是优美的，即虚拟机的操作系统将正确的关闭虚拟机。否则，系统在尝试优美地关闭后，将强制关闭虚拟机。这就像虚拟机突然掉电一样。

如果正在通过远程桌面终端监视虚拟机，则可以看到如所示的虚拟机关闭画面。

图 11-64 正在关闭的虚拟机

11.4.6 重启和重置虚拟机

虚拟机的“重启”和“重置”是一对容易混淆的名词。从本质上讲，功能都是重新启动

虚拟机。区别在于：“重启”相当于用户在 Linux 系统中调用“reboot”命令；而“重置”则相当于在服务器的机箱上触动“电源重置”按钮。

只有虚拟机的“运行状态”为“运行”时，重启和重置按钮才有效。如果你认为虚拟机应当在运行，而页面上显示的却是“关机”，请手动刷新页面获取最新的虚拟机状态。使用重启功能时，先要选择虚拟机，然后再点击“重启”按钮，在弹出的“是否确定重启”询问框中点击“确定”按钮后开始重启虚拟机，否则取消重启。重置操作和重启操作是类似的，这里不再描述。

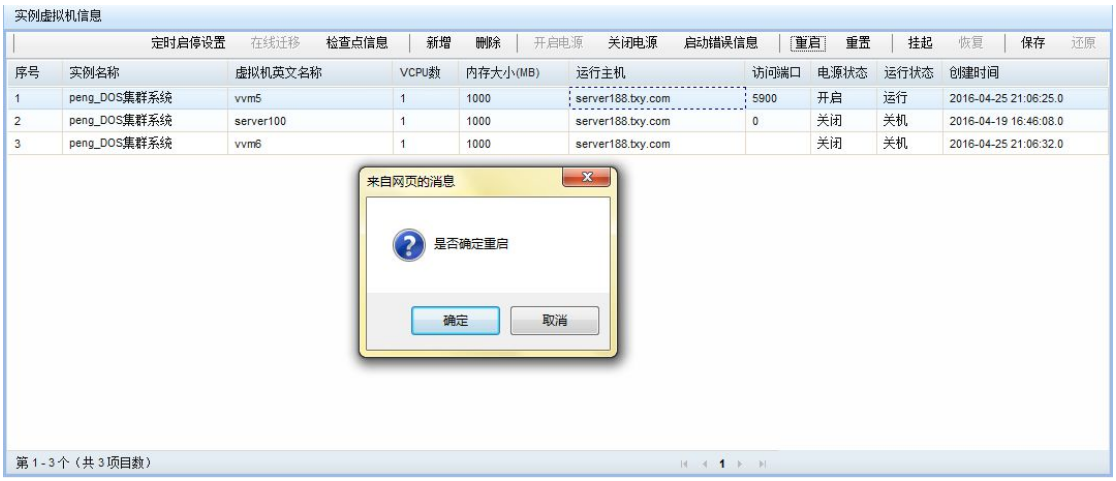


图 11- 46：重启确认画面

要注意的是：重启操作有可能不产生任何作用。原因是虚拟机操作系统忽略了外部的重启请求，因此操作被忽略。而重置操作无需虚拟机操作系统处理，因此总是会成功的，除非宿主机故障了。

11.4.7 挂起和恢复虚拟机

虚拟机的挂起和恢复相当于“暂停”和“继续”。当虚拟机挂起后，虚拟机上的系统就处于一种暂停的状态，不能进行任何的操作，恢复后则可继续运行。需要注意的是：挂起的虚拟机实际上仍然处于运行状态，因此处于挂起状态的虚拟机关闭后不能再恢复；而且挂起的虚拟机仍然占有系统资源。

如图 11-47 所示，虚拟机挂起前可操作，可以在操作系统登录界面的中输入字符（比如“aa”）。

图 11- 47：挂起前的虚拟机

选中当前运行的虚拟机，点击“挂起”按钮并确定，使虚拟机处于暂停状态。

图 11- 48：挂起操作的确认提示

如所示，虚拟机的运行状态由原来的“运行”转换成了“挂起”。

图 11- 49：成功挂起的虚拟机状态

接下来，再来对已经处于挂起状态的虚拟机进行输入操作（比如，想要在字符“aa”后面增加字符“bb”），发现不能完成。说明虚拟机挂起是系统处于暂停状态。

图 11- 50：挂起状态的虚拟机无法操作

现在选中已经挂起的虚拟机，如下图所示，点击“恢复”按钮并确定，使暂停的虚拟机恢复正常运行。

图 11- 51：恢复虚拟机的询问提示

这时，再来对已经恢复运行状态的虚拟机进行操作（比如，在字符“aa”后面增加字符“bb”），发现能够完成，如所示。说明虚拟机恢复运行状态。

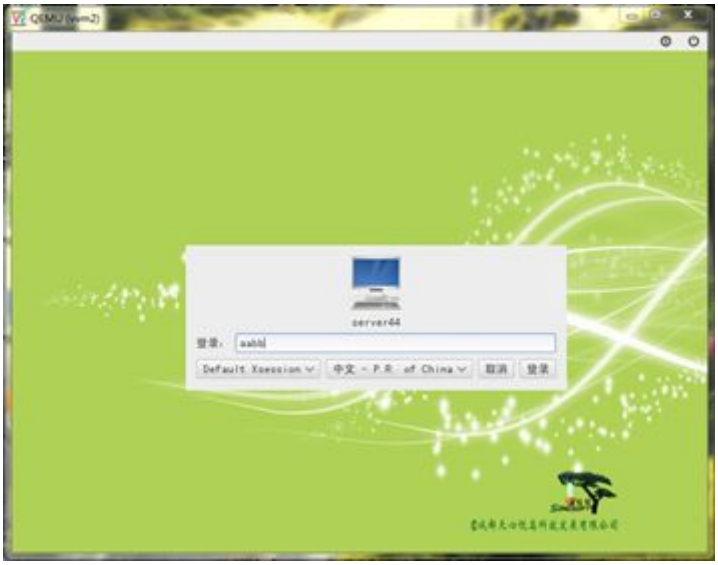


图 11- 52：恢复运行后的虚拟机

11.4.8 保存和还原虚拟机

这里将介绍虚拟机的保存和还原。虚拟机的保存过程是保存当前运行的虚拟机的内存到磁盘并关闭。还原的过程是使用保存的虚拟机内存文件来启动虚拟机。

图 11-53 所示是保存虚拟机的界面。保存虚拟机时，先选中想要保存的虚拟机，然后点击“保存”按钮，确定后开始保存。

图 11- 53：保存虚拟机的确认提示

保存成功后，虚拟机的运行状态由“运行”转变为“保存停止运行”，如图 11-54 所示。

图 11- 54：保存成功的虚拟机状态

选中处于“保存停止运行”状态的虚拟机，然后点击“还原”按钮，确定后开始还原，还原成功后虚拟机处于运行状态。

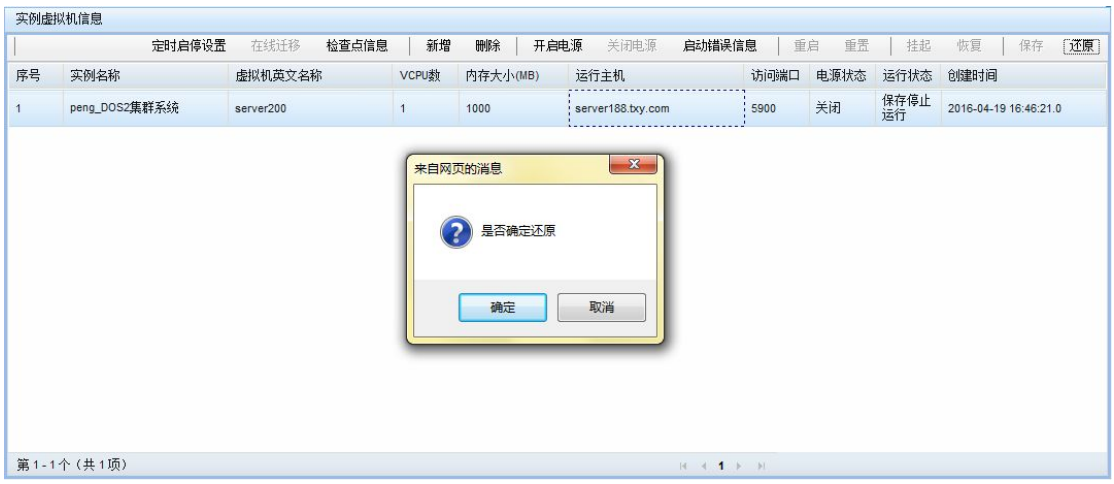


图 11- 55：还原保存的虚拟机

这时可以通过远程桌面终端操作正在运行的虚拟机。

11.5 虚拟机运行监控与查询

项目管理员登陆后，在导航栏上使用鼠标点击“运行监控与查询”，即可进入本小节即将介绍的用户界面。

本部分有两个内容，一个是动态信息实时监控，另一个是历史信息查询。由于虚拟机数量众多，加上更新周期较小，如果将运行记录都存储到数据库中则可能导致数据库负载过重。因此，虚拟机的动态负载信息以运行日志的方式保存在宿主机的本地文件系统中。一个项目可以同时运行多个虚拟机。用户指定要监控的虚拟机，一个页面同时监控的虚拟机数量不超过 3 个，并不断刷新显示的状态信息。

项目管理员也可以查询某个虚拟机实例的历史状态信息。这时需要选择虚拟机实例，然后输入查询的时间段。管理网站先查询虚拟机的历史事件信息表，找出指定时间段虚拟机实例运行的位置。然后请求宿主机返回虚拟机在这个时间段内的动态信息记录。管理网站将这些记录保存在内存中，然后以折线图的方式分别显示 CPU、内存、各个网络接口、各个块设备的负载情况。

虚拟机运行监控与查询页面的组织关系如下：

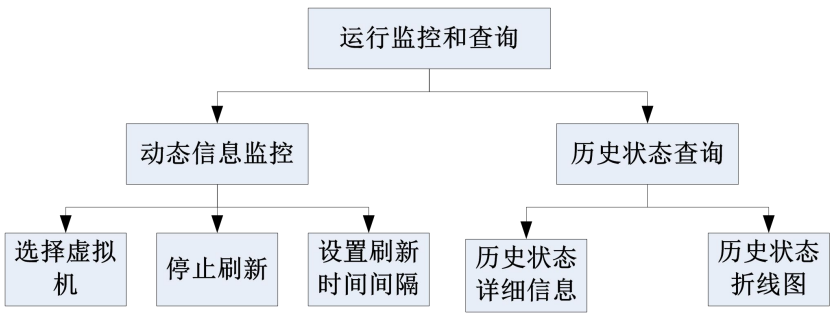


图 11- 56：虚拟机运行监控与查询

11.5.1 动态信息监控

下图为“动态信息监控”的界面。界面的工作区域分为横向的三块，预示着最多只能支持显示三个虚拟机的动态信息。

图 11- 57：动态信息监控操作界面

点击右上角的“选择虚拟机”按钮，系统会弹出“选择虚拟机”面板。在面板上有“虚拟机 1”、“虚拟机 2”和“虚拟机 3”三个选择输入框，均为可选的。如图 11-58 所示，这里同时选择了两个虚拟机。

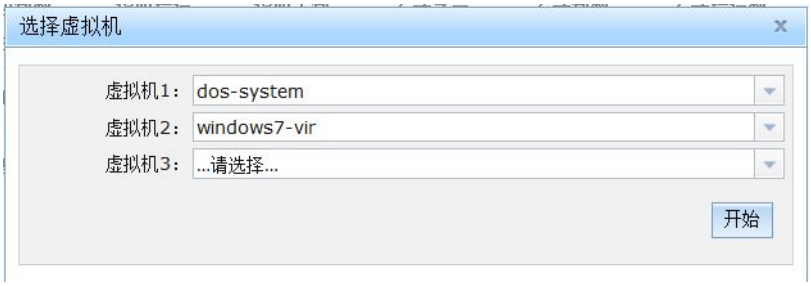


图 11- 58：选择需要实时监控的虚拟机

点击“开始”按钮，系统开始从运行服务器获取虚拟机的实时状态数据，并以下图的方式显示出来。显示的内容有：虚拟机、块设备和网络接口。“虚拟机”的内容总是会显示，如果不存在“块设备”或者“网络接口”，则不会显示。其中的虚拟机内容包括：虚拟机名称、运行状态、总内存、已用内存、CPU 数量、CPU 运行时间、运行主机和访问端口；其中的块设备内容包括：磁盘名称、总容量、已分配容量、实际占用容量、读取请求次数、读取字节数、写入请求次数、写入字节数、读取总时间和写入总时间；其中的网络接口内容包括：读取字节、读取包数、读取错误、读取丢包、发送字节、发送包数、发送错误数和发送丢包。图 11-57 为开始虚拟机实时监控后的运行截图。

虚拟机运行的实时数据是按照一定的时间间隔不断获取的。缺省的时间间隔为 10 秒钟。如果需要，可以根据需求将时间间隔进行调整。设置刷新时间可以完成这样的调整。时间间隔越小，刷新频率越快，系统资源消耗也越大。建议不要把时间间隔配置的太小，太小将消

耗服务器的系统资源。如下图所示，将缺省的时间间隔由原来的 10 秒钟修改为 30 秒钟，再点击右上角的“设置”按钮就能完成修改。

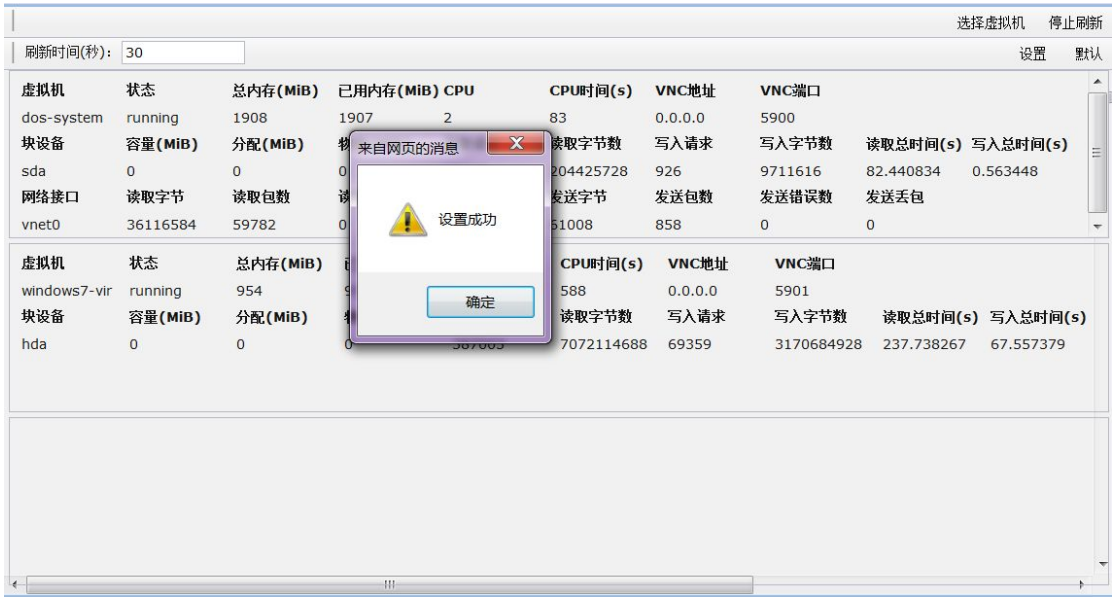


图 11- 59：修改实时监控的刷新时间

11.5.2 历史状态查询

历史状态在数字有机体虚拟机系统中分别以“数据”和“图表”两种方式展现出来。“数据”可供用户查询历史运行的细节；“图表”可供用户观察运行趋势，便于数据分析。“历史状态详细信息”就是把原始的“数据”提供给用户查看；“历史状态折线图”则是以“图表”的形式展现给用户。

图 11-60 为“历史状态查询”的主界面。它列出了所有的可以用作查询的历史状态的虚拟机。查看“历史状态详细信息”和“历史状态折线图”都需要先选择“虚拟机”。

图 11- 60：虚拟机状态查询中的虚拟机列表

以下描述获取“历史状态详细信息”的步骤。

第一步：选中想要查看的虚拟机记录，并点击“历史状态详细信息”按钮，弹出查询时

间选择面板；

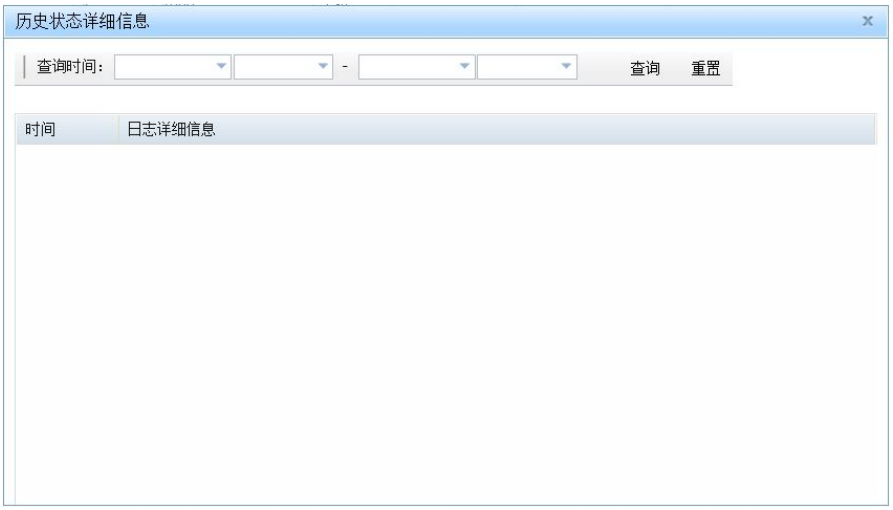


图 11- 61：输入历史状态查询的时间段

第二步：在“查询时间选择面板”中选择开始时间和结束时间，并点击“查询”按钮，就可以获取到指定时间内的历史运行数据。为了防止查询的时间段太大，记录数太多，每次查询返回的历史记录数量上限为 10000 行数据。历史运行数据的内容和实时监控虚拟机的内容是一致的，每条记录的内容有：虚拟机、块设备和网络接口。其中的虚拟机内容包括：虚拟机名称、运行状态、总内存、已用内存、CPU 数量、CPU 运行时间、运行主机和访问端口；其中的块设备内容包括：磁盘名称、总容量、已分配容量、实际占用容量、读取请求次数、读取字节数、写入请求次数、写入字节数、读取总时间和写入总时间；其中的网络接口内容包括：读取字节、读取包数、读取错误、读取丢包、发送字节、发送包数、发送错误数和发送丢包。

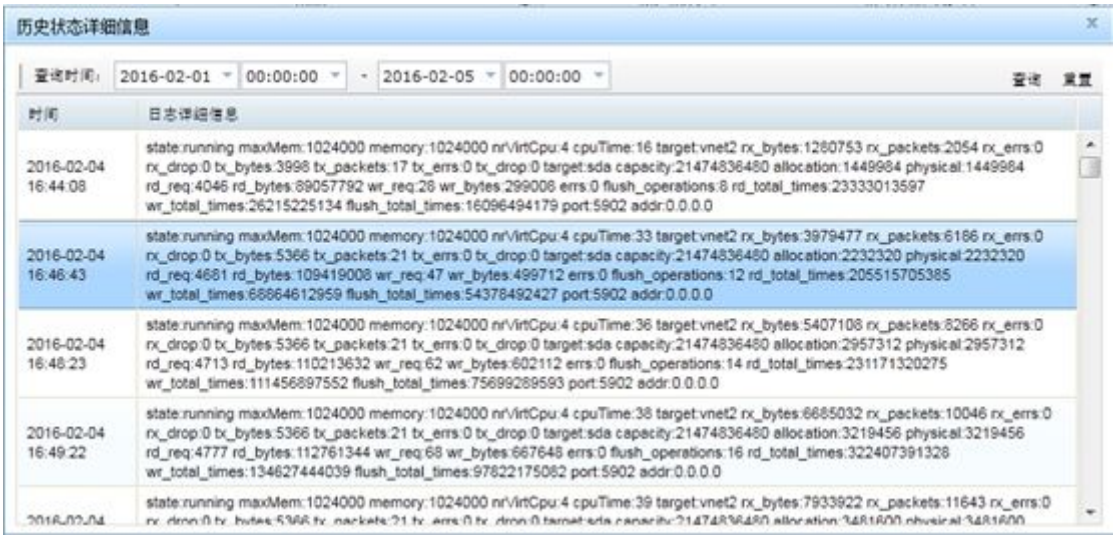


图 11- 62：查询到的历史状态信息

以下描述获取“历史状态折线图”的步骤。

第一步：选中想要查看的历史的虚拟机记录，并点击“历史状态折线图”按钮，弹出查询时间选择面板。该面板的和“历史状态详细信息”查询时的时间选择面板相同；

第二步：在“查询时间选择面板”中选择开始时间和结束时间，并点击“查询”按钮，就可以获取到指定时间内的历史运行数据，并根据这些数据生成折线图。统计折线图主要统计了虚拟机一段时间内的内存、磁盘 I0（读写数据）和网络 IO（收发数据）几个指标的使用趋势。如图 11-63 所示，分别采用不同颜色的折线来表示内存、网络等指标。

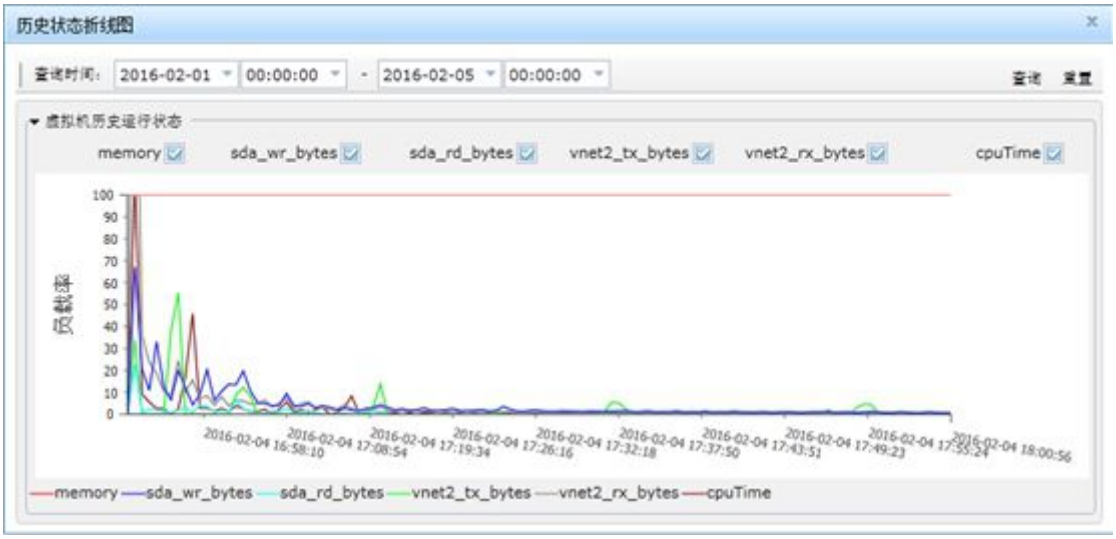


图 11- 63：查询到的历史状态折线图

如果用户不能很好地查看，可以反选各个指标，使它不显示，只查看想要的指标。例如，图 11-64 为只显示磁盘“sda”的写数据指标的历史记录折线图：

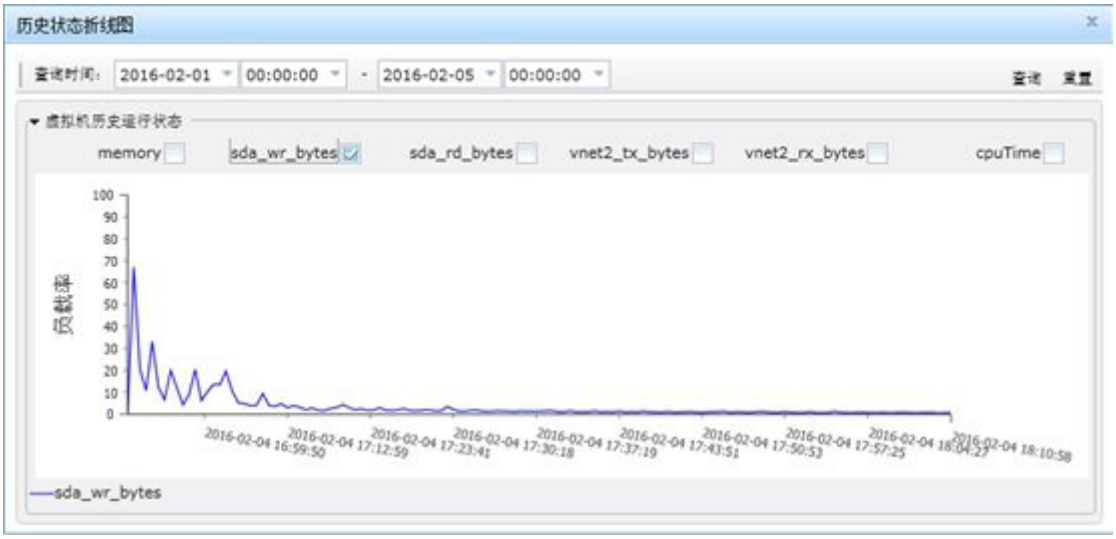


图 11- 64：磁盘写数据折线图

12. 虚拟机集群

12.1 概述

将一组虚拟机组合起来，共同为网络用户提供服务，这就形成了一个虚拟机集群。数字有机体的虚拟机集群具有负载均衡、故障屏蔽和动态虚拟机三大功能。和使用硬件设备部署的集群相比，前两个功能是共同具有的。而动态虚拟机则是数字有机体虚拟机系统特有的，通过动态增减虚拟机，可以更好的利用系统资源。另外，数字有机体系统可以自动的部署虚拟机集群，并屏蔽调度服务器的故障，这也是硬件负载均衡器难以实现的。

数字有机体虚拟机系统可以自动部署虚拟机集群的调度器。这就需要跟虚拟机集群配置相应的参数。每个虚拟机集群有名称、描述、服务提供者(管理员自己标记的，系统不处理)、默认权重、使用的虚拟网络、管理状态和绑定宿主机等信息。

必须为虚拟机集群指定一个虚拟网络，而且这个虚拟网络必须使用一个 IP 子网，即在网络管理中配置的子网。子网必须关联了一个地址池。地址池中的 IP 地址将被视作虚拟机的服务地址，即可以提供服务的地址。如果这些地址上的服务活动了，则自动加入到集群的可用服务器集中。而且，虚拟网络的类型不能是隔离网络和 macvtap 桥接网络，只能是 NAT、route、linux 桥接、ovs 桥接网络。

虚拟机集群的调度器可以由系统部署。部署调度的宿主机可以由管理员指定，即绑定在某台宿主机。如果未绑定宿主机，则由系统根据默认规则选择宿主机做调度器。默认权重是虚拟机参与服务时的默认权重。

每个虚拟机集群可由多个实例的虚拟机构成。这些虚拟机实例称为虚拟机集群的成员。这些虚拟机实例的某个网络接口必须使用集群关联的虚拟网络，从而其接口在同一个子网下。集群成员可按两种方式运行。一种是“管理员手动控制”，即由管理员自己启动或者关闭。这种成员称作静态成员，管理员可以指定其静态 IP 地址和调度权重，以便设定其参与服务的能力。另一种是“自动控制”，管理员配置实例的最少虚拟机数、最大虚拟机数、增加虚拟机的负载限制，减少虚拟机的负载限制和检查周期等参数。系统根据这些参数自动启动虚拟机。*要注意的时：虚拟机实例的所有虚拟机，无论他是系统自动启动的，还是管理员手动启动的，都被认为是集群的成员，集群成员的主机地址都应该属于使用的子网范围。*当一个虚拟机集群被启动时，系统将自动启动 “自动控制”的虚拟机，而“管理员手动控制”的虚拟机则由管理员控制。

一个虚拟机集群可以同时提供多个网络服务，因此可以设置多个 VIP（虚拟 IP）。每个 VIP 描述集群提供的一个网络服务。每个 VIP 包括协议类型、VIP、VPORT、最大连接数、是否会话持久和负载均衡方式。支持的协议类型为 TCP 和 UDP。会话持久方式现在只支持源地址方式，不支持 http cookie 和 app cookie，因此没有可选的。负载均衡方式包括轮转(wrr)、最小连接(wlc)和源地址哈希(sh)等。

可以为虚拟机集群的不同 VIP 配置不同的监控方式。每个监控方式的参数包括探测方式、检查端口（仅对 TCP 和 HTTP 有效）、探测间隔、超时时间(秒)、访问路径（仅对 HTTP

有效）和重试次数。支持的探测方式包括 PING、TCP 连接和 HTTP。PING 方式探测到的是整个虚拟机接口的死亡，这时不能设定 VIP，而是认为成员上的所有服务都无效。其他两种方式探测到的死亡可能是针对具体服务的，因此允许设定 VIP。对单个 VIP 来说，只能设定一种方式。而且，若设定了 PING 探测方式，则不再能设置任何其他方式。系统根据自动探测的结果设定参与服务的真实 IP。

本章将对虚拟机集群的配置作详细介绍，图 12-1 是虚拟机集群的页面结构图。

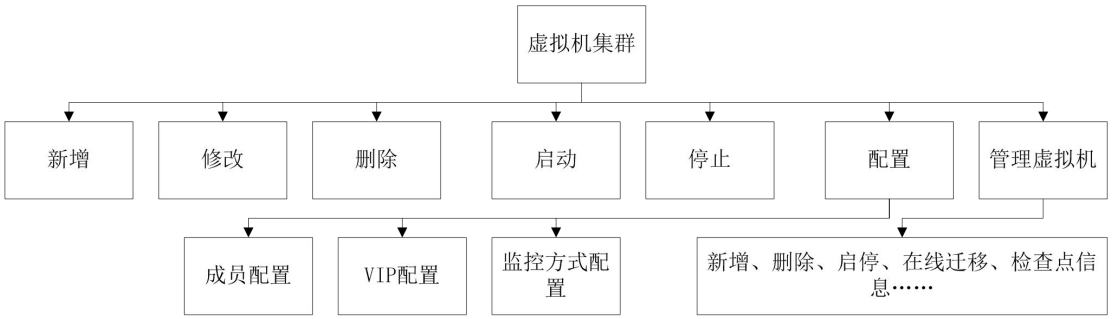


图 12-1：虚拟机集群的页面结构图

12.2 配置虚拟机集群

项目管理员登录后点击“虚拟机集群”即可进行虚拟机集群的配置。进入该部分的首页是“虚拟机集群列表”。这个页面以列表的形式列出所在项目的所有虚拟机集群信息，分页显示。页面工具栏的按钮由新增、修改、删除、启动、停止、配置和虚拟机管理组成。其次，还可以使用“启动错误信息”按钮来获取启动错误消息。查询工具栏可以按照虚拟机集群名称进行查询。列表的字段为：序号、名称、服务提供者、主机默认权重、管理状态、绑定调度服务器、当前调度服务器和描述。虚拟机集群的界面结构如图 12-2 所示。

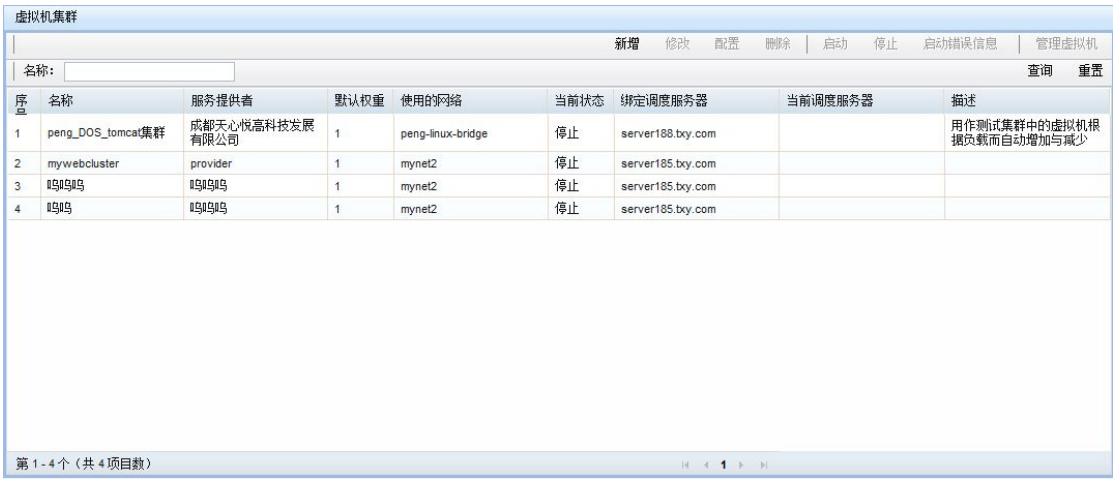


图 12-2：虚拟机集群列表

12.2.1 新增

点击 “新增”按钮弹出新建虚拟机集群面板（如图 12-3）。新增虚拟机集群时，仅输入集群的基本信息，即名称、描述、服务提供者、默认权重、使用的虚拟网络和绑定调度服

务器（非必选），新增集群成功之后默认是“关闭”状态。

名称、服务提供者、描述由管理员自己标记的，系统不处理。

集群必须指定一个虚拟网络，而且这个虚拟网络必须使用一个 IP 子网，即在网络管理中配置的子网。子网必须关联了一个地址池。网络的类型不能是隔离网络和 macvtap 桥接网络，只能是 NAT、Route、Linux 桥接、ovs 桥接网络。这里的下拉菜单只列出了符合要求的项目配置的虚拟网络。如果配置集群时需要手动配置“静态 IP”，该 IP 也应该属于子网地址池的范围。如果“静态 IP”没能按照要求配置，“监控配置方式”可能无效。而在配置 VIP 时，VIP 的地址最好不要在子网地址池的范围内。

默认权重是虚拟机参与调度时的服务权重。这是一个相对权重，范围在 1 到 255 间。建议取一个中间的数字。如果没有设定静态成员，则所有虚拟机的服务权重是相同的。权重数字越大表明虚拟机的服务能力越强。

如果未选择绑定宿主机，则由系统根据默认规则选定虚拟机集群的调度器。

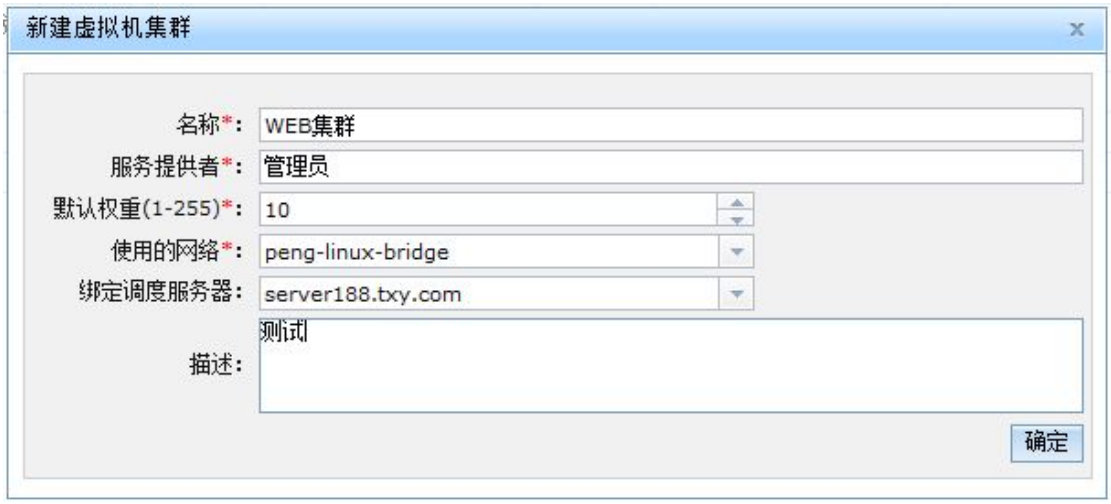


图 12- 3：新建虚拟机集群

12.2.2 修改

点击 “修改”按钮弹出修改虚拟机集群面板（如图 12-4）。修改虚拟机集群时，仅修改集群的基础信息，其他的修改在配置面板中。

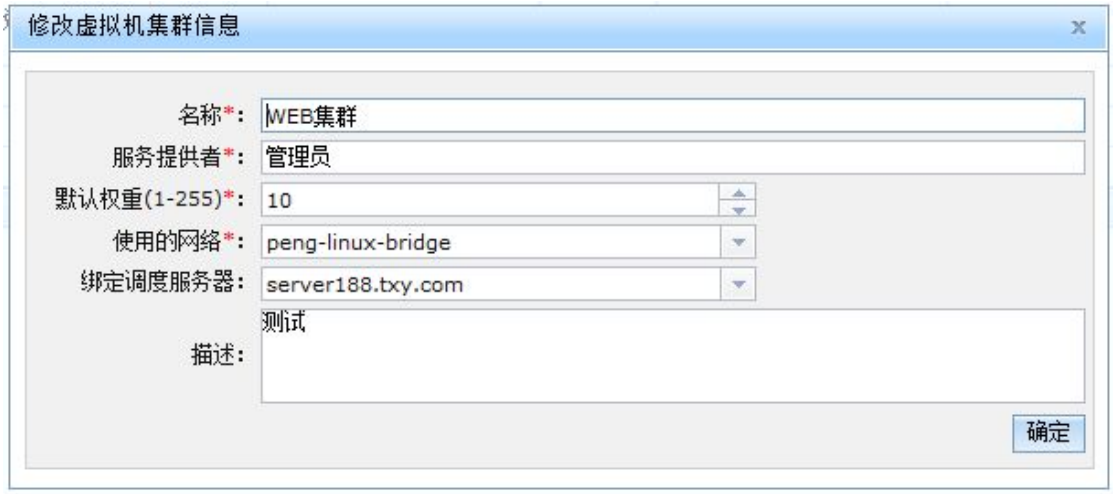


图 12- 4：修改虚拟机集群信息

12.2.3 删除

点击 “删除”按钮时，首先向用户确认，由用户点击确认后做相应操作，并返回操作结果。

12.2.4 启动、停止

项目管理员可以通过点击 “启动”、“停止”按钮来控制集群的启停状态。需要注意的是，配置了“静态 IP”的虚拟机需要手动启动。点击“停止”按钮时如果还有虚拟机正在运行，系统将做如图 12-5 的提示，点击“确定”按钮会关闭集群内所有虚拟机，点击“取消”按钮系统不会关闭虚拟机，但都会继续关闭集群。

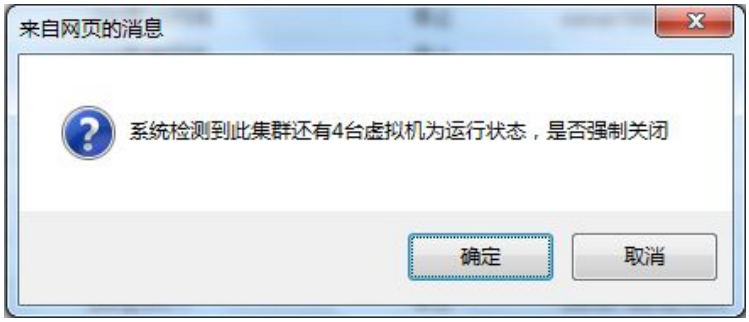


图 12- 5：关闭集群提示框

12.2.5 配置

新增虚拟机集群成功后点击“配置”按钮即可跳转到配置虚拟机集群页面，如图 12-6。在这里首先为集群配置成员，即提供服务的虚拟机实例；然后是配置集群的 VIP 信息，一个资源池可以配置多个 VIP；最后配置集群的监控方式，除“PING”方式外一个集群也可以配置多种监控方式。

图 12- 6：配置虚拟机集群

12.2.5.1 成员配置

点击“成员配置”表头切换到成员配置页面（如图 12-7）。页面上半部是集群成员列表，

下半部是对成员列表的操作，包括添加、删除、修改。

添加：先点击“清空”按钮清除数据，然后填写相关数据，点击“保存”按钮，等待操作结果。

修改：先选中成员列表的一条数据，然后修改，点击“保存”按钮，等待操作结果。

删除：先选中成员列表的一条数据，然后点击“删除”按钮，等待操作结果。

虚拟机实例应该先配好，这里只需选择即可。虚拟机的静态 IP 是可选的，当选中之后则使用权重加 IP 地址方式设置 IPVS 调度;否则后台监控程序将探测子网的地址池中的所有地址，发现有新地址可用时，以默认的权重加入到 IPVS 调度中。权重是虚拟机实例的服务权重。对每个非静态 IP 集群成员，还可以设定自动增减虚拟机参数。这些参数包括最大虚拟机数、最少虚拟机数、增加时的负载限制、减少时的负载限制、评估负载的时间间隔、增减虚拟机的时间间隔。当选择为静态 IP 时，系统不会动态的增减虚拟机数量，只会根据配置的 IP 去寻找虚拟机；相反为非静态 IP 时，系统将根据设置的负载上下限动态的增减虚拟机数量。

在选择虚拟机实例时应注意，选择的虚拟机实例配置的虚拟网络应与当前的集群配置的虚拟网络一致，否则不在一个子网内可能会发生相互通信不成功的情况。

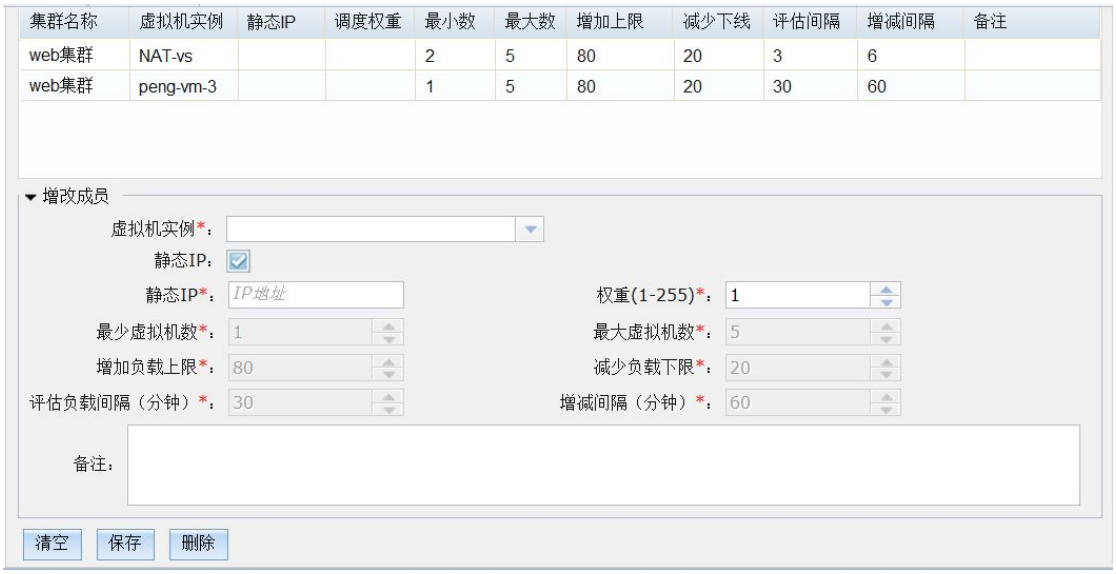


图 12- 7：成员配置页面

12.2.5.2 VIP 配置

点击“VIP 配置”表头切换到 VIP 配置页面（如图 12-8）。页面上半部是 VIP 列表，下半部是对 VIP 列表的操作，包括添加、删除、修改。

添加：先点击“清空”按钮清除数据，然后填写相关数据，点击“保存”按钮，等待操作结果。

修改：先选中 VIP 列表的一条数据，然后修改，点击“保存”按钮，等待操作结果。

删除：先选中 VIP 列表的一条数据，然后点击“删除”按钮，等待操作结果。

一个虚拟机集群可以有多个 VIP，每个 VIP 包括协议类型、是否会话持久、客户地址掩码、超时时间、虚拟 IP、虚拟端口、最大连接数、转发方式和负载均衡方式。支持的协议

类型为 TCP、UDP。会话持久方式目前只支持源地址方式，暂不支持 http cookie 和 app cookie。当选中会话保持时，需要填写客户地址掩码（客户地址比较的掩码）和超时时间。虚拟 IP 即虚拟服务的单一 IP 入口地址，配置时需要和使用的子网同网段，最好不要在子网的地址池范围内。虚拟端口一般情况下与服务的端口一致。如果虚拟服务端口和真实服务的端口不同，则需要外部网络设备来完成映射，例如反向 NAT 服务器。系统需要为每个 VIP 启动一个 IPVS 虚拟服务。

配置 VIP 时应注意：

负载均衡方式包括权重轮转(wrr)、最小连接(wlc)、从不排队（nq）、最短期望延迟(sed)、轮转(rr)、最少连接数(lc)、局部最少连接数(lblc)。

转发方式（向真实服务器转发请求的方式）包括 Direct-Route 和 NAT。

NAT：是一种最简单的方式，所有交互数据必须通过均衡器；

Direct-Route：直接路由，调度器和真实服务器必须在同一个网段，通过修改 IP 包的 MAC 地址进行转发。

图 12- 8：VIP 配置

12.2.5.3 监控方式配置

点击“监控方式配置”表头切换到监控方式配置页面（如图 12-9）。，页面上半部是监控方式列表，下半部是对监控方式列表的操作，包括添加、删除、修改。

添加：先点击“清空”按钮清除数据，然后填写相关数据，点击“保存”按钮，等待操作结果。

修改：先选中监控方式列表的一条数据，然后修改，点击“保存”按钮，等待操作结果。

删除：先选中监控方式列表的一条数据，然后点击“删除”按钮，等待操作结果。

可以为资源池的不同 VIP 配置不同的监控方式。监控方式目前支持三种方式，即 PING、TCP 和 HTTP。每个监控方式的参数包括探测方式、探测端口（仅对 TCP 和 HTTP 有效）、探测间隔（秒）、超时时间（秒）、重试次数、对应的 VIP（仅对 TCP 和 HTTP 有效）和访

问路径（仅对 HTTP 有效）。

PING 方式探测到的是整个虚拟机接口的死亡，这时不能设定 VIP，而是认为成员上的所有服务都无效，并且若采用该方式就不能再配置其他方式。TCP 和 HTTP 两种方式探测到的死亡是针对具体服务的，因此允许设定 VIP，并且可以重复配置或者配置多种方案。需要注意的是：使用 HTTP 方式时，将使用“http://被探测 IP:检查端口/访问路径”来请求服务，因此它必须为可访问路径。例如 http://192.168.2.190:8080/demo/demo.jsp 的访问路径为 demo/demo.jsp，访问端口为 8080，被探测的 IP 为 192.168.2.190。被探测的 IP 是虚拟网络的地址池中的 IP。

设定的所有监控方式将同时有效。IPVS 调度服务器同时采用设定的方式监控每个活动的虚拟机，以便有效地，准确的，及时的发现故障的虚拟机。发现虚拟机或者虚拟机上的服务故障时，将从 IPVS 调度中去除它。

图 12- 9：监控方式配置

12.2.6 虚拟机管理

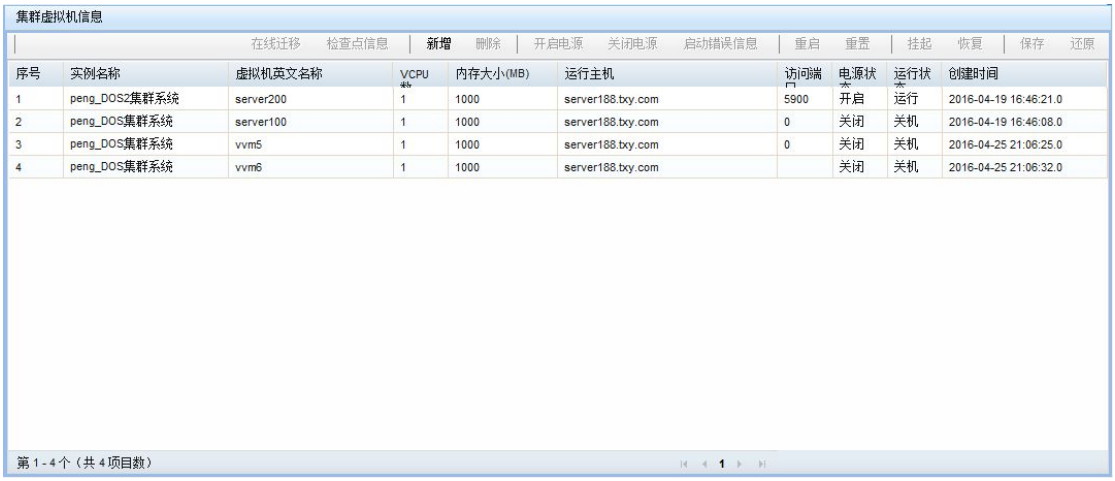


图 12- 10：集群虚拟机列表

点击 “管理虚拟机”按钮即切换到集群虚拟机信息页面（如图 12-10）。这里可以为集群内的实例新增、删除、关闭、开启、保存、还原等虚拟机操作。此页面与“虚拟机管理”页面大体一致，不同在于这里显示的是多个实例的虚拟机信息，所以这里不再对此页面作详细的介绍。集群的虚拟机可以手动新增，系统也会根据配置自动增减，手动新增时需要选择实例名称。

12.3 部署虚拟机集群

在部署虚拟机集群之前最好关闭所有节点的 iptables 和 selinux 服务，即 service iptables stop 和 setenforce 0。

数字有机体虚拟系统利用 Linux 的 IPVS 模块实现 IP 请求的转发。它同时支持两种请求转发方式，即 NAT 和 Direct-Route （直接路由）。

虚拟机集群的服务节点，即虚拟机，都在同一个子网内。这是因为系统只监视一个虚拟网络内的服务，而这个虚拟网络只有一个子网。虚拟机集群的调度节点也有一个 IP 子网内的地址，即子网的网关。因此，部署 NAT 和直接路由方式的虚拟机集群是最合适的方式。

12.3.1 NAT

12.3.1.1 概述

NAT 模式是 IPVS 最常用的一种模式。相比于 Direct-Route 模式，NAT 模式更容易部署，且不需要为虚拟机配置公网 IP 地址，从而可以节约部署成本。

采用 NAT 转发方式时，虚拟机的服务 IP 可采用私有地址。虚拟机的默认网关即子网的网关。子网的网关担任调度节点。系统自动部署网关和调度服务，因此很容易配置。

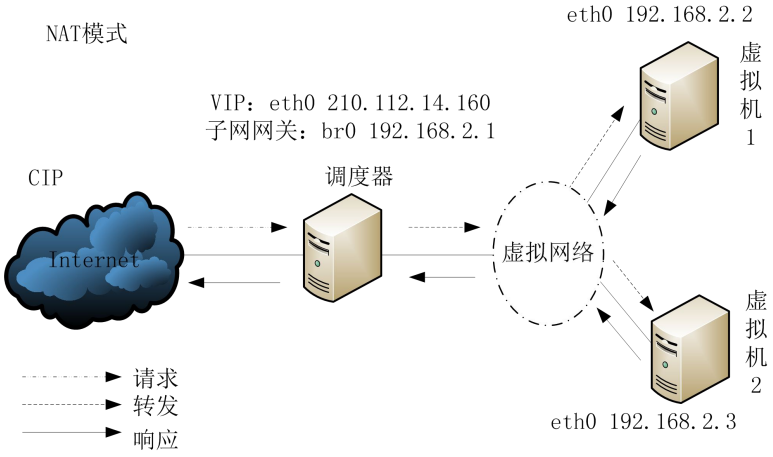


图 12- 11：NAT 模式

采用 NAT 转发的消息处理过程如图 12-11 所示。客户的请求使用服务的 VIP 地址（公网地址）作为目标地址，即 210.112.14.160。Internet 将请求路由至子网网关所在的节点，也是服务的调度节点，也是 VIP 地址的绑定节点。调度节点选择一台虚拟机作为真实服务节点，并登记客户地址(即源地址)，客户端口（即源端口）、真实服务节点的地址和真实服务端口等信息于内核中。然后用真实服务节点的 IP 替换请求包中的目的地址，用网关的 IP 替换请求包中的源地址，并将请求转发给虚拟机。虚拟机以为这是一个来自网关的请求。因此请求处理的响应包也发送给网关。网关在收到响应时，查找内核中登记的映射记录，将响应包的目的 IP 再替换为客户地址，目的端口替换为客户端口，源 IP 替换为 VIP。响应请求包通过互联网出口发送，从而被正确返回给客户。

12.3.1.2 部署实例

下面以配置 web 服务的虚拟机集群为例来详细说明 NAT 模式集群的配置步骤。dos_vs_test/sessions.jsp 网页是用来测试 session 保持的一个测试页面，如图 12-12 所示（非 IPVS 模式）。页面记录了一个访问地址和一个真实服务器地址。这里的访问地址就是 vip，真实服务器地址就是虚拟机的地址。

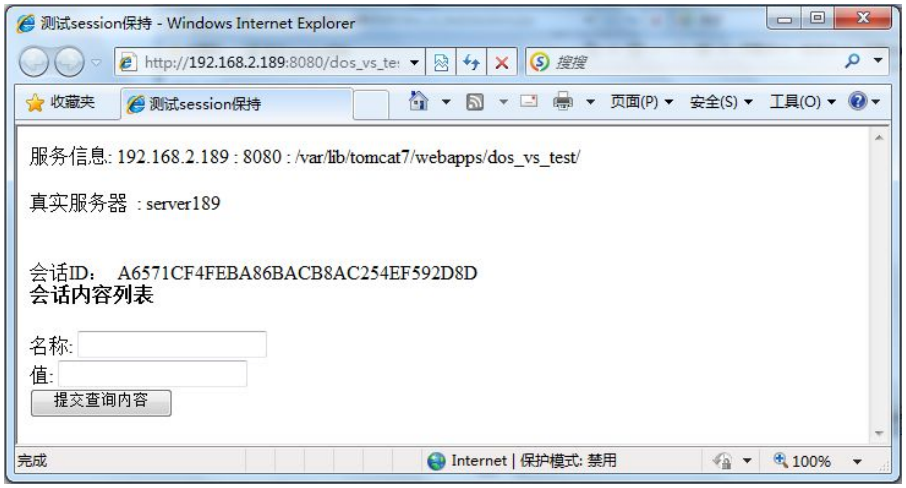


图 12- 12：测试 session 保持页面

因环境限制，现假设“192.168.2.x”网段公网网络，用 192.168.2.6 作为公网服务的

IP，即 WEB 服务的 VIP。虚拟机部署在 192.168.7.x 子网上，可以认为它是一个私有的子网。每台宿主机有两个网卡。网卡 eth0 用于连接公网，有公网 IP 地址，即 192.168.2.x 网段的地址。网卡 eth1 用于部署虚拟机的网络，采用 Linux 桥接方式，以便跨宿主机部署虚拟机。

1）在系统中先建立 192.168.7.x 的子网，配置如图 12-13 所示。

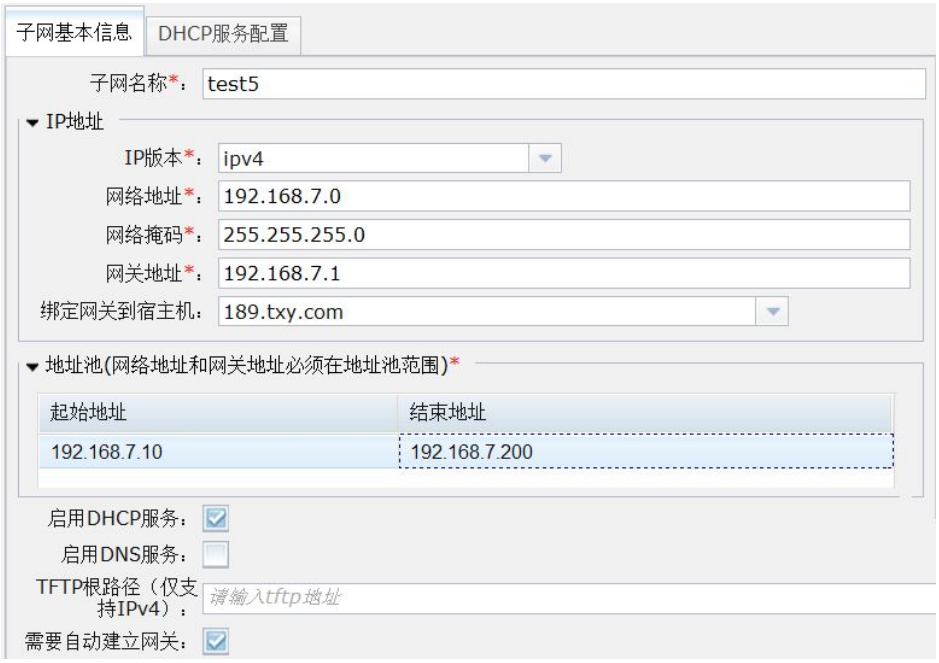


图 12- 13：子网 test5 的配置

这里将子网的网关设为 192.168.7.1，并将网关绑定到“189.txy.com”，使得集群调度器也将绑定在该宿主机上。地址池的地址为 10 到 200，10 以前的留在内部服务使用。启用 DHCP 服务，以便虚拟机可以自动分配到服务地址。由系统自动建立网关，以便部署网关和集群调度器。DHCP 服务没有进一步配置，使用默认设置。

2）建立桥接虚拟网络，以便跨宿主机部署虚拟机集群。桥接网络的配置如图 12-14 所示。网络的其他参数对集群部署没有影响。关键是这个网络使用子网 test5，即用于部署虚拟机集群的子网。转发网络接口使用宿主机的 eth1，即用于部署虚拟网络的接口。转发延迟最好不设置，这里仅是用于测试的。网络本身不能绑定到宿主机，因为它要跨多个宿主机部署。

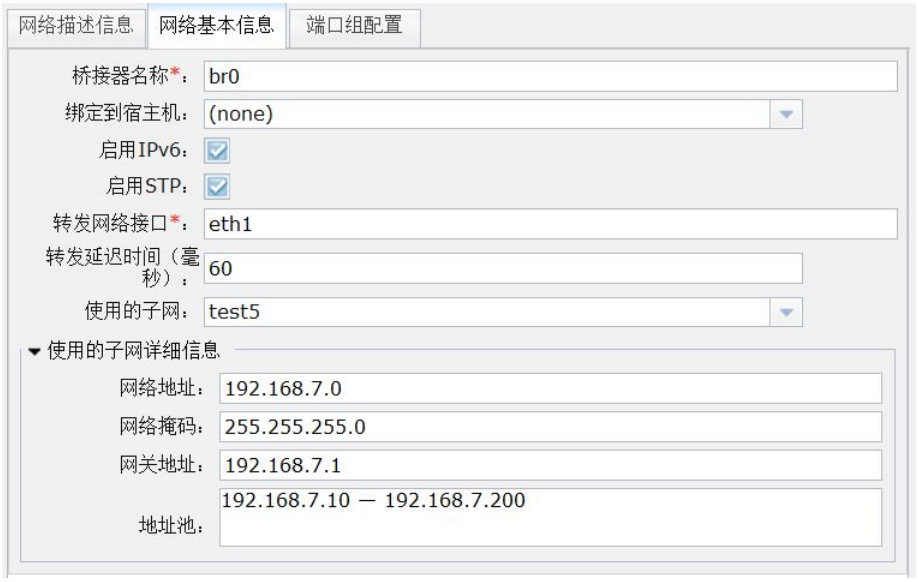


图 12- 14：集群桥接网络的配置

3）建立虚拟机实例。这个过程不再描述，它包括配置虚拟机运行参数，在虚拟机中安装操作系统，部署网站。其中关键的是虚拟机的网络接口 eth0 使用 DHCP 方式配置。虚拟机启动后自动启动 WEB 服务。WEB 服务的端口为 8080。

4)新建虚拟机集群。虚拟机集群的配置如图 12-15 所示。

图 12- 15：NAT 方式的虚拟机集群

该虚拟机集群采用上面配置的虚拟网络，没有绑定宿主机，由系统按照规则选择。由于子网网关绑定了宿主机，因此虚拟机集群的调度器也绑定在该宿主机上。

5）配置虚拟机集群的成员。即用前面配置的虚拟实例作为虚拟的成员。配置如图 12-16 所示。

图 12- 16：NAT 集群的成员配置

这里的虚拟机成员采用自动控制方式，在集群启动时即自动启动至少一个虚拟机。在负载超过 80%时则增加虚拟机。

6）配置虚拟集群的 VIP，即虚拟服务。其配置如所示。

图 12- 17：NAT 集群的 VIP 的配置

由于 WEB 服务使用的协议是 TCP，因此这里协议类型为 TCP。需要测试会话保持，所有这里配置了会话保持。客户地址掩码为全 255，即每个客户地址独立保持。转发方式设定为 NAT。负载均衡方式可以选择，这里使用默认的。

图 12- 18：NAT 集群中的监控方式实例

7）配置虚拟机集群的监控方式。在配置了 VIP 后，必须配置监控方式，否则 IPVS 调度无法自动配置。这里采用 HTTP 方式监控 WEB 服务，如图 12-18 所示。这里的关键点是采用 HTTP 监控方式时，访问路径一定要配置为网站的访问路径，检查端口即 WEB 服务端口。。

8）配置完集群后就可以启动集群。在“虚拟机集群”页面选择一个集群，点击“启动”按钮即可启动虚拟机集群，并且弹出执行结果对话框。

9）查看虚拟机集群运行情况。

在虚拟机集群浏览页面上，可以查看虚拟机集群的运行宿主机。现在可以看到当前的运行宿主机为“189.txy.com”。登录这台服务器，在命令终端可以用 ifconfig 查看服务器网络接口的配置情况。其中可以看到虚拟网络使用的桥接器 br0 的配置。由于它同时是子网的网关，因此网关地址绑定在 br0 上。在集群调度服务器上，服务的虚拟 IP 都绑定在桥接器的别名上，这里使用网络号做别名编号，因此是 br0:38。

```
root@server189:~# ifconfig
br0       Link encap:Ethernet  HWaddr 00:1e:67:21:6c:41
          inet addr:192.168.7.1  Bcast:192.168.7.255  Mask:255.255.255.0
          inet6 addr: fe80::21e:67ff:fe21:6c41/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:434 errors:0 dropped:0 overruns:0 frame:0
          TX packets:5 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:194388 (189.8 KiB)  TX bytes:398 (398.0 B)

br0:38    Link encap:Ethernet  HWaddr 00:1e:67:21:6c:41
          inet addr:192.168.2.6  Bcast:192.168.2.255  Mask:255.255.255.255
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
```

图 12- 19：NAT 集群的调度服务器的接口配置

在虚拟机集群的调度服务器上，可以通过 ipvsadm 命令查看 IPVS 的配置情况，如图 12-20 所示。这时可看到已经配置的 VIP 以及可以使用的虚拟机的服务地址。

```
root@server189:/var/log# ipvsadm -L -n
IP Virtual Server version 1.2.1 (size=4096)
Prot LocalAddress:Port Scheduler Flags
  -> RemoteAddress:Port           Forward Weight ActiveConn InActConn
TCP  192.168.2.6:8080 nq persistent 600
  -> 192.168.7.38:8080            Masq    10     0          0
```

图 12- 20：ipvsadm 命令

10）验证虚拟机集群提供的 WEB 服务。此时再通过 VIP 访问页面，真实服务器 IP 会出现虚拟机的 IP，如图 12-21。从返回结果可以看出，真实服务的节点时 server145，这是虚拟机的名称。不过，服务信息中显示的是“192.168.7.38”，而不是请求的 vip 地址。这恰好是 NAT 处理的结果。在转发请求给虚拟机时，调度服务器已经将请求包中的目的地址替换为虚拟机的 IP 地址。如果启动多台虚拟机，则来自不同客户的情况可以在这些虚拟机间分别处理。

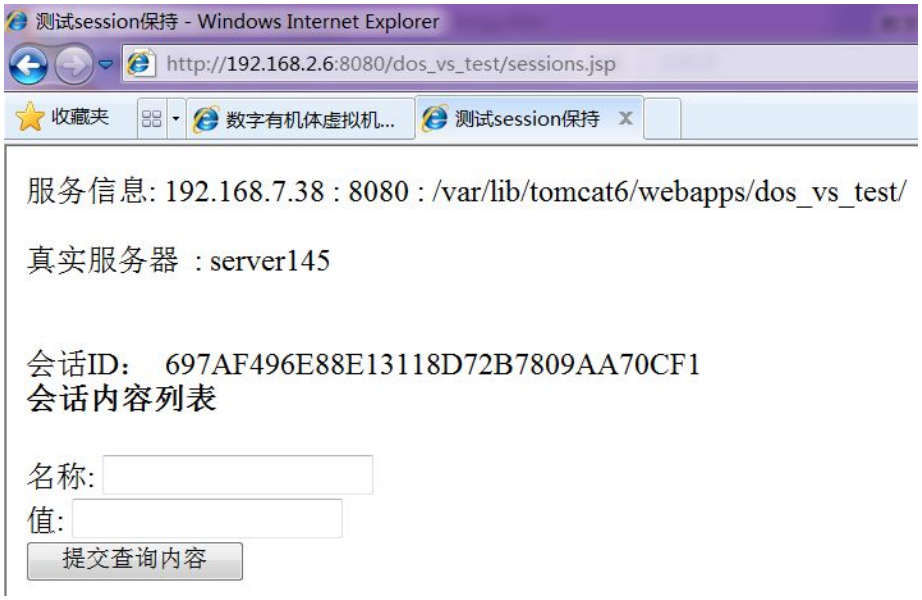


图 12- 21：通过 VIP 访问到虚拟机

12.3.2 Direct-Route（直接路由）

12.3.2.1 概述

和 NAT 转发模式相比，Direct-Route 避免了 NAT 模型的每个数据包都要经过负载均衡节点的情况。Direct-Route 模型只有请求的时候才会经过负载均衡节点，回应的数据包由虚拟机直接响应用户不需要经过负载均衡节点。但是，为了使响应消息能够回到客户节点，就需要虚拟机也能直接和外部网络通信，这就需要虚拟机有公网 IP 地址。这是直接路由方式应用的一个限制。

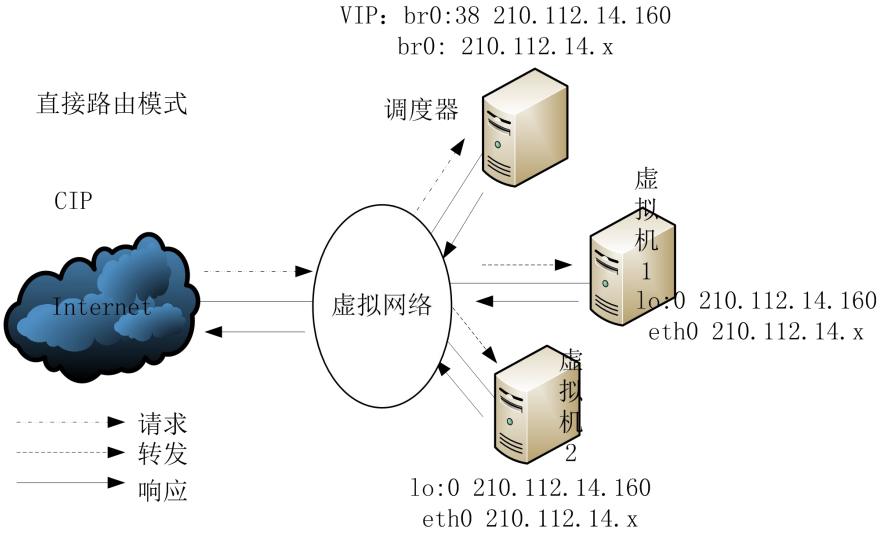


图 12- 22：直接路由模式

和 NAT 部署不同的是，由于每台虚拟都需要有公网地址，为了方便，大多数时候给每台虚拟机配置和 VIP 相同网段的公网地址，如图 12-22 所示。客户访问服务时，请求的目标地址仍然是 VIP。系统将 VIP 绑定在集群调度器的桥接器上，这样请求将有互联网路由到

调度器。调度器按照规则选择一台虚拟机处理请求。这是调度查询虚拟机的 eth0 的 MAC 地址，将请求用虚拟机的 MAC 转发出。由于虚拟机和调度在同一个交换网内，因此虚拟机将收到请求。虚拟机处理请求的响应直接发送给客户（这里请求的源地址没有被修改）。

部署直接路由模式的虚拟机集群有以下几点需要注意：

1）虚拟网络必须采用桥接网络，且运行虚拟网络的宿主机接口必须在同一个交换网中。

2）每个虚拟机都要配置公网地址，且 VIP 需要手动配置在 lo:0 上。

3）集群的调度节点不一定是虚拟网络的网关。

12.3.2.2 部署实例

以 NAT 部署的环境为例，采用直接路由时，虚拟机也要使用公网的 IP 地址，因此虚拟网络的子网不能是上面 192.168.7.x，必须是 192.168.2.x。为此另外配置一个子网，其配置如图 12-23 所示。

图 12- 23：直接路由集群的子网配置

这个子网由外部路由器担任网关，因此无需系统部署网关。DNS 和 DHCP 服务也由外部服务器提供，因此也不启动。

配置虚拟网络的方式和 NAT 集群的网络配置相似，只是将子网选择为这里配置的公网服务子网。

配置虚拟机实例的方式也和 NAT 集群相似，只是虚拟机的网络接口才使用新配置的网络。每个虚拟机内也要绑定服务 IP 地址，不过该地址要求绑定在回环设备 lo 上，作为 lo 的别名。例如 ifconfig lo:0 VIP/32。注意，绑定 IP 地址时的掩码为 255.255.255.255。默认回环设备上的地址不会用于响应 ARP 请求，因此绑定的 IP 地址仅仅用于虚拟机接受负载均衡节点转发来的用户请求。

配置虚拟机集群的方式和配置 NAT 集群相同，只是转发方式改为直接路由。

在启动该虚拟机集群后，系统将自动选择调度节点。可以在虚拟机集群列表中看到当前的调度节点。登录调度节点可以查看网络接口的使用情况。如图 12-24 所示。和 NAT 集群的

示例不同，这里集群的调度节点并不担任网关，因此 br0 上没有配置网关地址。服务的 VIP 仍然绑定在桥接器的别名上。

```
root@server189:/var/log# ifconfig
br0       Link encap:Ethernet  HWaddr 00:1e:67:21:6c:41
          inet6 addr: fe80::21e:67ff:fe21:6c41/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:1489 errors:0 dropped:0 overruns:0 frame:0
          TX packets:6 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:887546 (866.7 KiB)  TX bytes:468 (468.0 B)

br0:38    Link encap:Ethernet  HWaddr 00:1e:67:21:6c:41
          inet addr:192.168.2.6  Bcast:192.168.2.255  Mask:255.255.255.255
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
```

图 12- 24：直接路由集群示例启动后调度节点接口配置

现在用“ipvsadm -l -n”来查看 IPVS 的配置情况，如图 12-25 所示。现在虚拟机的服务 IP 也是 192.168.2.x 网段的。其中的 forward 方式显示为“Route”，即直接路由方式。

```
root@server189:/var/log# ipvsadm -l -n
IP Virtual Server version 1.2.1 (size=4096)
Prot LocalAddress:Port Scheduler Flags
  -> RemoteAddress:Port           Forward Weight ActiveConn InActConn
TCP  192.168.2.6:8080 nq persistent 600
  -> 192.168.2.33:8080            Route   10     0          0
```

图 12- 25：直接路由集群示例启动后的 IPVS 配置

在虚拟机内，网络接口的配置如图 12-26 所示。其中 eth0 是用 DHCP 方式自动配置的，不过提供服务是物理路由器。lo:0 是手动配置的，绑定了 VIP 地址。

```
root@server145:~# ifconfig
eth0      Link encap:Ethernet  HWaddr 52:54:00:50:ed:8b
          inet addr:192.168.2.33  Bcast:192.168.2.255  Mask:255.255.255.0
          inet6 addr: fe80::5054:ff:fe50:ed8b/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:9176 errors:0 dropped:0 overruns:0 frame:0
          TX packets:140 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:442760 (432.3 KiB)  TX bytes:13262 (12.9 KiB)
          Interrupt:10 Base address:0xc000

lo        Link encap:Local Loopback
          inet addr:127.0.0.1  Mask:255.0.0.0
          inet6 addr: ::1/128 Scope:Host
          UP LOOPBACK RUNNING  MTU:16436  Metric:1
          RX packets:1740 errors:0 dropped:0 overruns:0 frame:0
          TX packets:1740 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:104428 (101.9 KiB)  TX bytes:104428 (101.9 KiB)

lo:0      Link encap:Local Loopback
          inet addr:192.168.2.6  Mask:0.0.0.0
          UP LOOPBACK RUNNING  MTU:16436  Metric:1
```

图 12- 26：直接路由集群的虚拟机接口配置

现在，可以在客户浏览器中用 vip 请求 WEB 服务，图 12-27 是浏览结果。现在，服务仍然是有虚拟机提供的。不过，服务信息中显示的是“192.168.2.6”，不再是虚拟机 eth0 网

卡的地址。这是因为直接路由的请求是直接转发给虚拟机的，虚拟机内也绑定了 VIP，因此服务信息仍然是 VIP。

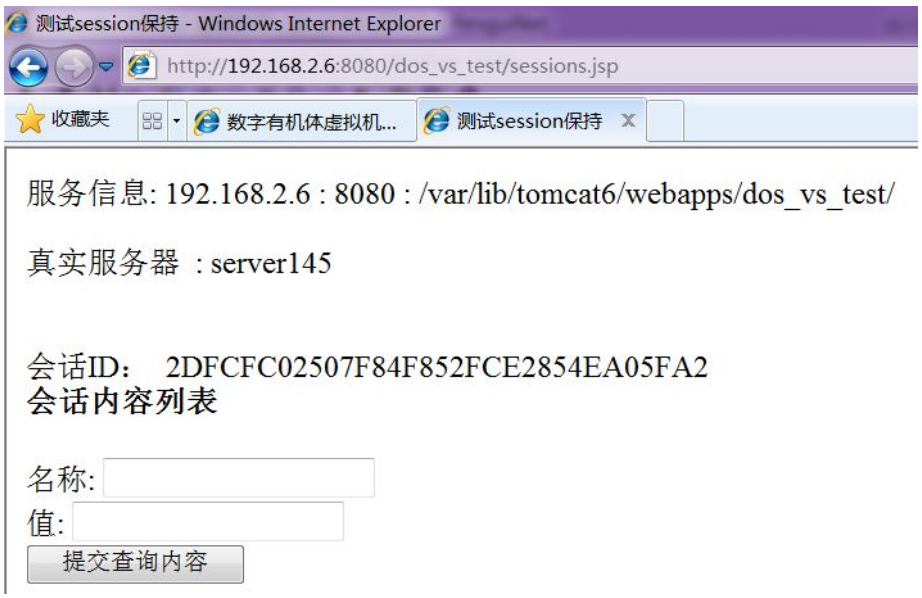


图 12- 27：直接路由集群的服务效果

12.4 运行管理

在管理系统新增配置好虚拟机集群后，默认是开启状态，也可以手动点击按钮来控制集群的启停操作。

启动成功后，可以通过“ipvsadm –L”命令查看其运行状况，如图 12-28 所示。

```
IP Virtual Server version 1.2.1 (size=4096)
Prot LocalAddress:Port Scheduler Flags
  -> RemoteAddress:Port           Forward Weight ActiveConn InActConn
TCP  192.168.3.2:http-alt nq persistent 600
  -> 192.168.3.159:http-alt       Route   10     0          0
```

图 12- 28：ipvsadm -L

宿主机的虚拟 IP 不需要手动创建。但是，如果采用的是直接路由转发方式，在虚拟机内部需要手动创建一个虚拟回环地址，例如 ifconfig lo:0 VIP netmask 255.255.255.255。

13. 其他管理功能

13.1 操作日志审计

操作日志仅系统管理员可以查看。用系统管理员用户名登录系统，点击“日志管理”面板，即展示用户操作日志，如图 13-1 所示。通过操作日志审计功能，系统管理员可以随时了解整个系统的操作运行情况，及时发现系统异常事件及非法访问行为。

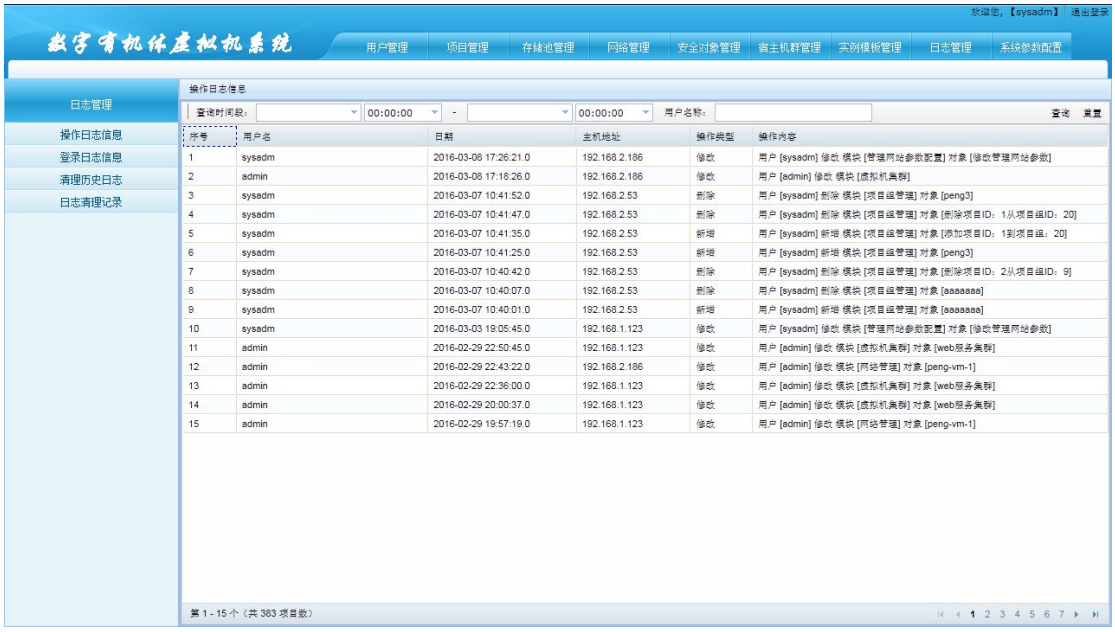


图 13-1：日志管理

13.1.1 操作日志信息

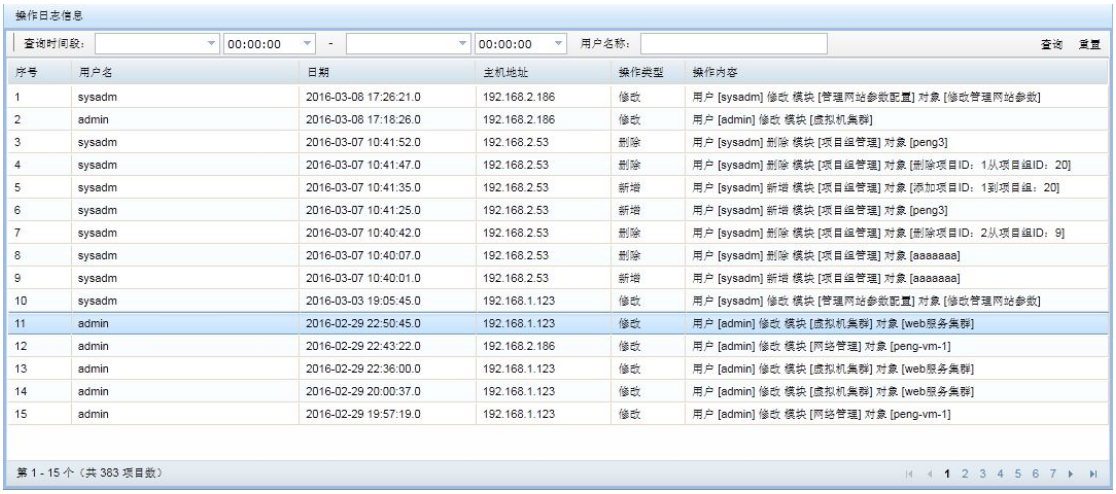


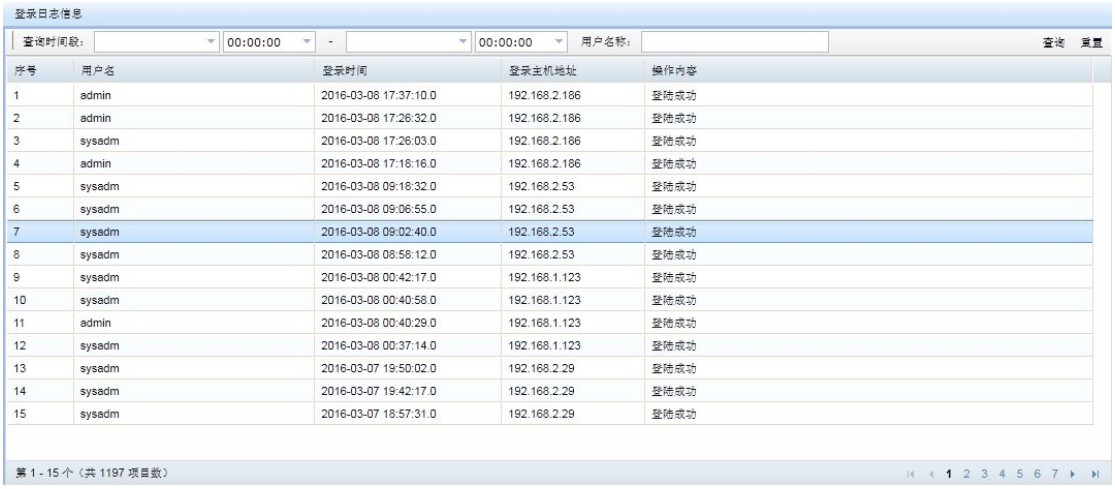
图 13- 2：操作日志内容

点击右侧栏“操作日志信息”即可显示出所有的用户操作日志，如图 13-2 所示，分页显示。显示的内容包括用户名、日期、主机地址、操作类型和操作类容。用户名是进行该操作的用户的账号。主机地址是操作请求发出的主机 IP，即用户使用的客户机的 IP 地址。操

作类型目前包括新增、修改和删除。操作内容记录了某个用户在某个模块做了什么事情。

系统提供了两种过滤功能。一是以时间方式过滤，二是以用户名方式过滤。时间过滤方式必须输入一个完整的时间段。用户名方式是模糊匹配。亦可两种方式同时匹配。然后点击“查询”按钮即可。

13.1.2 登录日志信息



13- 3：登录日志信息

点击右侧栏“登录日志信息”即可显示出所有的用户登录日志，如图 13-3：登录日志信息所示，分页显示。显示的内容包括用户名、登录时间、登录主机地址、操作类容。用户名是进行该操作的用户的账号。主机地址是操作请求发出的主机 IP，即用户使用的客户机的 IP 地址。操作内容只记录了成功的操作。

系统提供了两种过滤功能。一是以时间方式过滤，二是以用户名方式过滤。时间过滤方式必须输入一个完整的时间段。用户名方式是模糊匹配。亦可两种方式同时匹配。然后点击“查询”按钮即可。

13.1.3 清理历史日志

图 13- 4：清理历史操作日志

点击右侧栏“清理历史日志”即可显示出清理历史操作日志和登录页面，如图 13- 4。选择一个时间，点击“清理”按钮即可清除所选时间之前的所有日志，包括操作日志和登录日志，并将这一清理日志事件记录到日志清理历史信息页面内。

13.1.4 清理操作日志记录

本模块的主要功能是记录管理员清理操作日志和登录日志的事件，不能删除。点击左侧栏“日志清理记录”即可显示出所有日志清理的历史信息，分页显示，如图 13-5 所示。

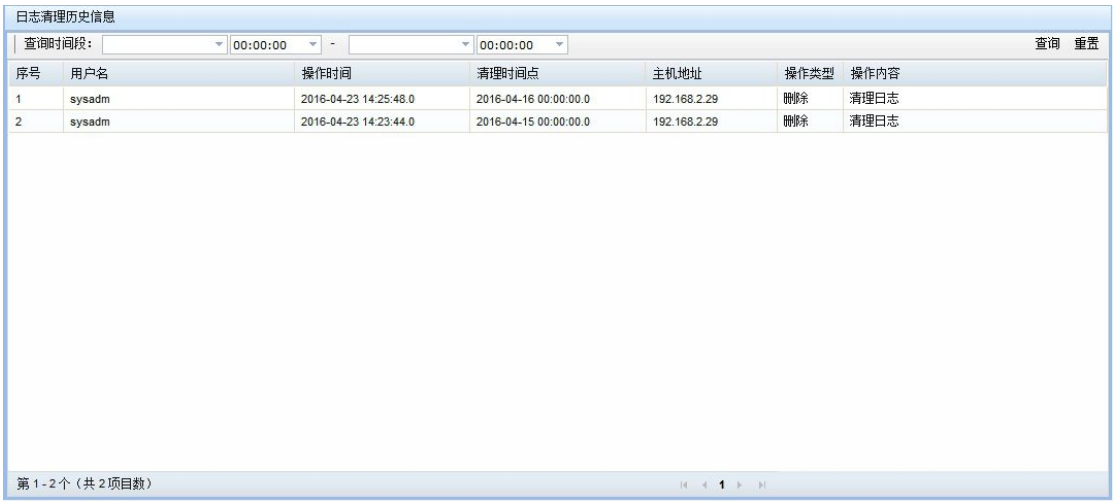


图 13-5：清理日志记录

显示的内容包括用户名、操作时间、清理时间点、主机地址、操作类型和操作类容。用户名是进行该操作的用户的账号。主机地址是进行操作的主机 IP，即系统管理员使用的客户机。操作时间是发生清理事件的时间。清理时间点是指所清理的时间点（时间点之前的所有日志都被清除）。

13.2 系统参数配置

本模块的主要功能是方便系统管理员配置后台系统（dosvmd）、管理网站（dosvm）的运行参数。在系统管理员登录后，点击“系统参数配置”面板，即可打开系统参数配置页面，如图 13-6 所示。

图 13- 6：系统参数配置

13.2.1 系统运行参数配置

点击右侧栏“系统运行参数配置”即可显示出系统运行参数配置页面，如图 13-7 所示。下面分别说明各项的用途：

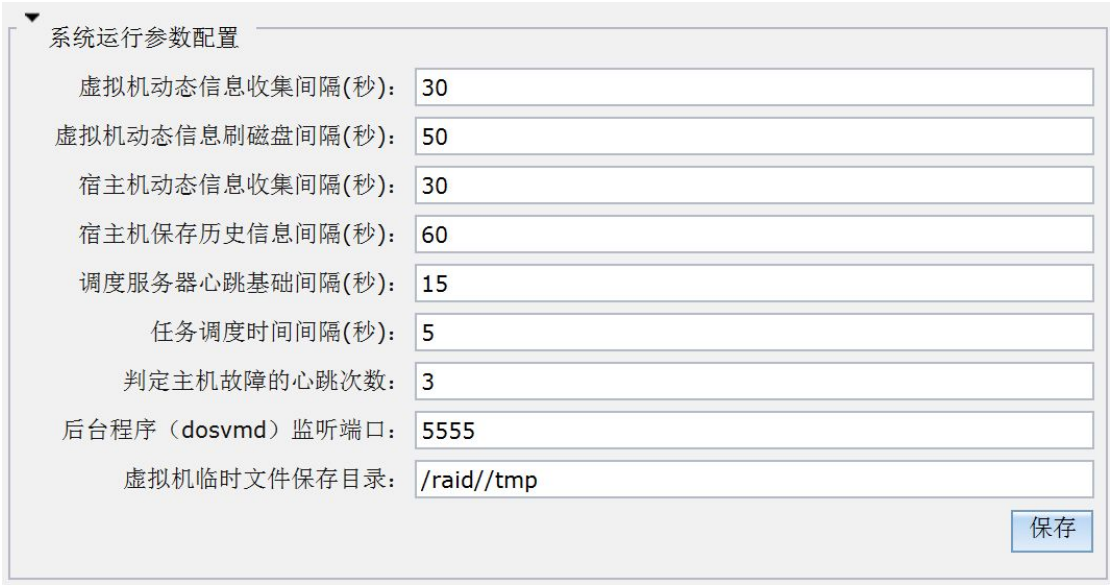


图 13- 7：系统运行参数配置

1) 虚拟机动态信息收集间隔：后台 dosvmd 程序用来获取虚拟机动态信息的间隔时间，单位秒，2～1800 之间；
2) 虚拟机动态信息刷磁盘间隔：后台 dosvmd 程序获取虚拟机动态信息之后刷磁盘的间隔时间，单位秒，4～1800 之间，必须大于虚拟机动态信息收集间隔；
3) 宿主机动态信息收集间隔：后台 dosvmd 程序用来获取宿主机动态信息的间隔时间，单位秒， 5～1800 之间；
4) 宿主机保存历史信息间隔：后台 dosvmd 程序用来保存宿主机历史信息的间隔时间，单位秒，10～3600 之间，该值必须大于宿主机动态信息收集间隔；
5) 调度服务器心跳基础间隔：后台 dosvmd 程序调度服务器心跳基础间隔时间，单位秒，2～300 之间；
6) 任务调度时间间隔：后台 dosvmd 程序调度工作间隔，单位秒，5～300 之间；
7) 判定主机故障的心跳次数：后台 dosvmd 程序判死主机前，允许主机丢失的心跳次数，1～10 次；
8) 后台程序（dosvmd）监听端口：后台 dosvmd 程序的监听端口，主要监听管理网站的请求，1～65535 之间，最好大于 1024；
9) 虚拟机临时文件保存目录：虚拟机产生的临时文件的保存路径。

13.2.2 管理网站参数配置

点击右侧栏“管理网站参数配置”即可显示出管理网站参数配置页面，如图 13-8。其中 SSL 相关的设置请参考 4.4.6.2 章节。下面分别说明各参数的用途：

1) 后台程序（dosvmd）主机地址：后台程序（dosvmd）所在的 IP 地址。一般而言，部署管理网站的本机如果同时也部署了 dosvmd 程序，这里的“主机地址”配置为本机 IP 地址。如果本机仅仅部署了管理网站，则需要将“主机地址”设定为运行有 dosvmd 程序的主机。仅需配置一个主机即可。
2) SSL 是否使用：管理网站与后台程序（dosvmd）通信时是否使用 SSL 通信加密，该配置必须要和 dosvmd 一致，两者要么都使用 SSL 加密，要么都不使用。

3） SSL 服务器证书路径： PKCS12 格式的服务器证书的路径，例如：第 4 章制作的证书 server-cert.p12。配置为全路径名。

4） SSL 根证书路径： PKCS12 格式的根证书路径，例如：第 4 章制作的证书 cacert.p12。配置为全路径名。

5） SSL 密码：输入 SSL 证书的密码。

6） 重复 SSL 密码：重复输入 SSL 证书的密码。

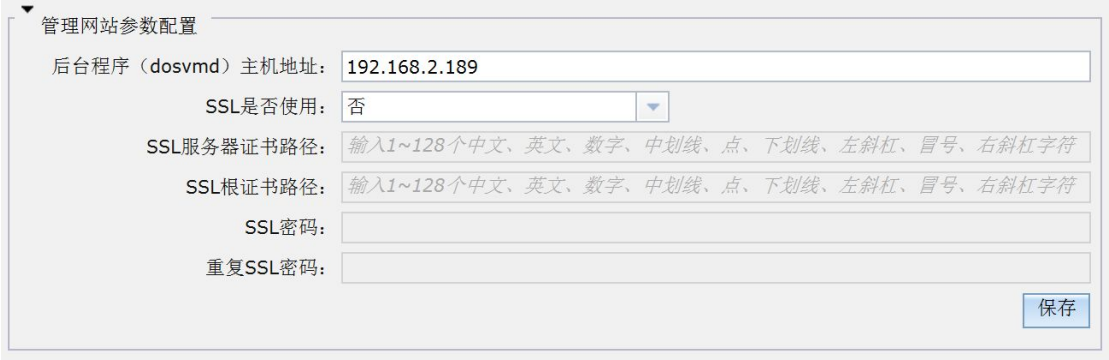


图 13- 8：管理网站参数配置

14.出错处理

如果在使用本产品时，您发现了某些缺陷与不足，或是需要帮助，请即时致函我们。我们的电子邮箱是：tianxinyue@126.com。

要想获取最新的在线用户手册和了解最新的相关信息，您可访问我们的公司 Web 站点 www.tianxinyue.com。