

ZettaStor 产品 用户手册

鹏云网络
2014 年 11 月 10 日

目 录

| | |
|--------------------------|----|
| 第一章 概述..... | 1 |
| 第二章 系统登录与退出..... | 3 |
| 2.1 系统登录..... | 3 |
| 2.2 系统退出..... | 3 |
| 第三章 存储管理..... | 4 |
| 3.1 域管理..... | 4 |
| 3.1.1 新建域..... | 4 |
| 3.1.2 修改域..... | 5 |
| 3.1.3 删除域..... | 6 |
| 3.1.4 查询域..... | 6 |
| 3.2 磁盘管理..... | 6 |
| 3.2.1 卸载和挂载磁盘..... | 6 |
| 3.3 存储池管理..... | 6 |
| 3.3.1 新建存储池..... | 7 |
| 3.3.2 修改存储池..... | 7 |
| 3.3.3 删除存储池..... | 8 |
| 3.3.4 查询存储池..... | 8 |
| 3.4 卷管理..... | 8 |
| 3.4.1 管理卷界面..... | 9 |
| 3.4.2 创建卷..... | 9 |
| 3.4.3 配置卷访问规则..... | 11 |
| 3.4.4 配置卷缓存类型..... | 12 |
| 3.4.5 配置卷 QOS..... | 12 |
| 3.4.6 挂载驱动和连接、断开客户端..... | 14 |
| 3.4.7 卸载驱动..... | 22 |
| 3.4.8 查询卷..... | 22 |
| 3.4.9 卷详细信息..... | 23 |

| | |
|-------------------------|----|
| 3.4.10 删除卷..... | 23 |
| 3.4.11 回收卷..... | 23 |
| 3.4.12 扩展卷..... | 23 |
| 3.4.13 迁移卷..... | 24 |
| 3.4.14 卷的克隆..... | 24 |
| 3.4.15 卷的快照..... | 25 |
| 3.4.16 孤儿卷..... | 26 |
| 第四章 系统管理与监控..... | 26 |
| 4.1 主页..... | 26 |
| 4.2 系统服务管理..... | 28 |
| 4.2.1 服务管理界面..... | 28 |
| 4.2.2 服务管理操作..... | 29 |
| 4.3 系统性能监控与管理..... | 29 |
| 4.3.1 创建性能任务..... | 29 |
| 4.3.2 启动性能任务..... | 30 |
| 4.3.3 停止性能任务..... | 31 |
| 4.3.4 删除性能任务..... | 31 |
| 4.3.5 利用性能任务监控卷性能..... | 31 |
| 4.3.6 利用性能任务监控实例性能..... | 32 |
| 4.4 数据重构与再平衡..... | 33 |
| 4.4.1 数据重构..... | 33 |
| 4.4.2 数据负载均衡..... | 33 |
| 4.5 系统告警监控与管理..... | 33 |
| 4.5.1 查询告警..... | 33 |
| 4.5.2 移除告警..... | 34 |
| 4.5.3 告警模板管理..... | 34 |
| 4.5.4 通过邮件发送告警..... | 35 |
| 4.6 用户管理..... | 36 |
| 4.7 许可证管理..... | 38 |

| 编制人 | 联系电话 | 时间 |
|-----|-------------|------------|
| 扈宇春 | 13381056202 | 2015-6-23 |
| 陆娜 | 13404128236 | 2016-10-24 |
| | | |
| | | |

第一章 概述

ZettaStor 分布式块设备存储系统（简称 ZettaStor DBS）是软件定义的分布式存储。它运用分布式计算技术把大量标准 x86 服务器的存储介质进行聚合，将这些存储资源整合成为既具备传统 SAN / NAS 的企业级功能和特性，又具有高弹性、高扩展性、高可靠性的存储系统，可称做 Server SAN。ZettaStor 系统可以与 OpenStack、Hadoop、VMware 和 fusion computer 进行无缝对接。

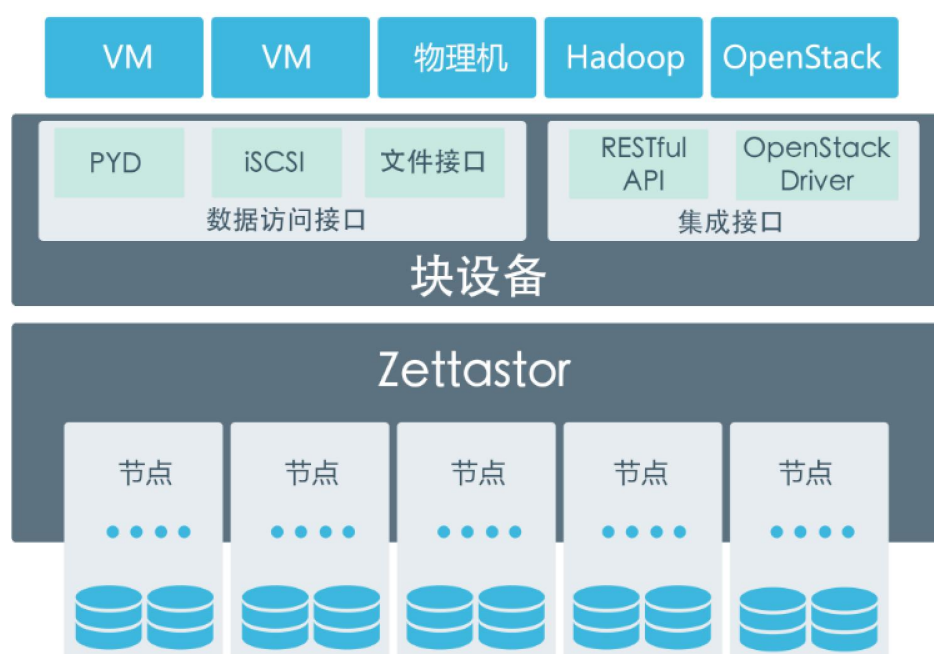


图 1 框架图

ZettaStor DBS 直接管理磁盘裸设备，无需由文件层转化，效率更高、性能更好。

ZettaStor DBS 通过 PYD 和 iSCSI 为客户机提供块存储服务。

ZettaStor DBS 可以为客户机提供快照、克隆、精简配置、Qos 保障、备份与恢复、数据复制等存储服务。并通过数据重新平衡(rebalance)、数据重构(rebuild)、性能加速等功能，为系统提供高可靠、高性能保证。

ZettaStor DBS 包括 InfoCenter、ControlCenter、MonitorCenter、DriverContainer、DataNode、DIH、Console、deployment_daemon 八个模块：

- ◆ deployment_daemon 是其他各模块通信的基础;
- ◆ DIH 监控管理所有服务的服务状态;
- ◆ InfoCenter 是系统的信息中心, 管理系统的配置信息(元数据信息)、性能和告警数据。元数据信息包括系统中所有的 Volume 信息、DataNode 信息、账户信息。一个 Volume 由多个 Segment 组成, Segment 分布在不同的 DataNode 上。Volume 信息包括 Volume 所有的 Segment 信息和 Volume 的挂载信息。其中 Segment 信息中包含 Segment 到 DataNode 的映射关系。DataNode 信息包含 DataNode 的存储资源, 具体为每个 DataNode 总共有多少空间、已使用多少空间、还有多少空间可以使用;
- ◆ ControlCenter 是系统的控制中心, 对外提供 Restful 管理接口。通过此接口, 客户可以发送创建 Volume、挂载 Volume、删除 Volume、罗列多个 Volume 信息、获取单个 Volume 信息等类型的请求到系统, 然后 ControlCenter 调用相应的模块完成相应的处理;
- ◆ MonitorCenter 是系统的告警模块, 针对系统出现的各类异常情况给出相应的警告以便于及时处理和维护;
- ◆ DriverContainer 为系统的网络驱动容器, 管理系统的所有网络驱动。当 DriverContainer 从外部接收挂载 Volume 的请求时, 便从容器选择一个驱动同 Volume 挂载起来供客户机使用。目前支持业内流行的 SCSI 和 iSCSI 协议;
- ◆ DataNode 为存储模块, 一个存储节点上部署一个 DataNode。DataNode 管理存储节点上的所有用于存储的磁盘, 并接收网络驱动发过来的读写请求, 进行硬盘读写操作。DataNode 集群采用 P2P 协议, 各个 DataNode 上保存各自的元数据信息, 并向其他的 DataNode 通报, 因此无需元数据中央节点;
- ◆ Console 系统的管理中心, 提供 Web 界面供用户管理整个系统, 包括账户管理、保护域管理、存储池、卷管理、节点管理、磁盘管理、告警和性能监控、统计分析、快照、克隆、服务管理等。

第二章 系统登录与退出

2.1 系统登录

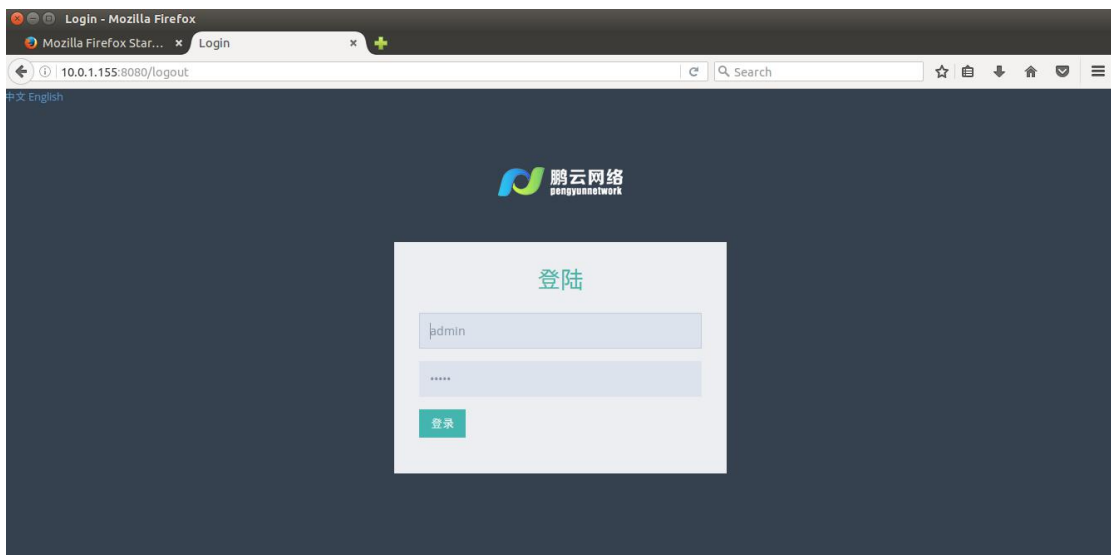
1、系统 Console 模块访问地址为：

操作：<http://【Console 的 IP 地址】:8080>

默认用户名：admin 密码：admin

admin 用户具有系统的最高权限，请及时在账户管理界面修改密码。

2、进入登陆页面，如图所示：

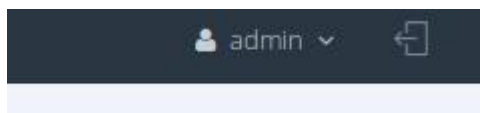


3、语言切换。

登陆页面左上角有“中文”和“English”两个链接，单击分别切换到相应的操作界面。

2.2 系统退出

安全退出登陆账号，直接单击最右边图标，或者鼠标放到用户上并选中 Log Out，如图所示。



第三章 存储管理

ZettaStor DBS 以块设备的形式向外提供存储服务，Operator(操作者)通过 Console 创建好 Volume(卷)以后，然后在界面上将 Volume 通过 DriverContainer 挂载到网络驱动上去，系统提供的网络驱动有 ISCSI 和 PYD。用户可以根据 DriverContainer 返回的挂载信息，在某个客户端主机通过网络驱动客户端连上挂载的网络驱动，最后便可以在客户端主机对挂载的 Volume 进行读写操作了。

日常使用 ZettaStor DBS，我们通常先新建域-->修改域-增减 DataNode 节点-->管理存储池-增减磁盘-->新建卷-->挂载驱动-->读写卷-->卸载驱动，给卷建立快照，回滚快照，扩展卷，等等很多功能。

3.1 域管理

域管理将 DataNode 节点划分为不同的范围，此范围可以根据 DataNode 的磁盘性能进行划分，也可以按不同的地域，不同的管理权限进行划分。创建域后，在创建 volume 时选择创建的域，该卷的 segment 将平均分配到域包含的 DataNode 节点之中，不会扩散到所有的 DataNode 节点。

域管理支持域的创建、删除、域信息查看、添加或者删除存储节点等操作。并可以选择存储节点使用哪些存储介质做主存(持久化存储)和缓存。

3.1.1 新建域



The image shows a '新建域' (New Domain) dialog box. It has a title bar with the text '新建域' and a close button (X). The main area contains two input fields: '域名称' (Domain Name) and '详细信息' (Detailed Information). Below these fields are two buttons: '创建' (Create) in blue and '取消' (Cancel) in red.

域名称：域的名称，唯一标识，不能修改，不能重复；

详细信息：对该域的描述，可写可不写。

操作：单击“创建”后，系统弹出域创建成功的信息。

3.1.2 修改域

操作：对域中所包含的 DataNode 节点进行增减，包括增加节点和移除节点。

- 增加节点：此节点纳入存储系统，做为存储节点提供存储服务。

选中某个域，单击“修改”，系统跳到修改域界面，如图所示。

| 实例ID | 所在组编号 | 主机IP地址 | 端口 | 当前状态 |
|---------------------|-------|------------|-------|------|
| 753591236041939957 | 0 | 10.0.1.159 | 10011 | OK |
| 3792445768498654111 | 3 | 10.0.1.156 | 10011 | OK |
| 1521360080721762850 | 2 | 10.0.1.158 | 10011 | OK |
| 160914925193676965 | 1 | 10.0.1.157 | 10011 | OK |

选择未被使用的 DATANODE 节点，单击“修改”，就完成了增加节点的操作。

操作：单击域名旁边的“+”，可看到如下信息：

| | 域名称 | 描述信息 | 当前状态 | 剩余容量 | 操作 |
|--------------------------|--------|------|------|---------|-------|
| <input type="checkbox"/> | domain | | 可用 | 120(GB) | 管理存储池 |

| 实例名称 | 实例ID | 所在组编号 | 主机IP地址 | 端口 | 当前状态 | 管理 |
|----------|---------------------|-------|------------|-------|------|------------|
| DataNode | 69945187670361856 | 1 | 10.0.1.159 | 10011 | OK | 移除Datanode |
| DataNode | 2068714593061261035 | 3 | 10.0.1.157 | 10011 | OK | 移除Datanode |
| DataNode | 3756104078996979529 | 0 | 10.0.1.158 | 10011 | OK | 移除Datanode |
| DataNode | 4533402302526432155 | 2 | 10.0.1.156 | 10011 | OK | 移除Datanode |

当前状态：当前域的状态分为可用和不可用；

剩余容量：代表该域中包含的所有 Datanode 节点的剩余容量；

- 移除节点：此节点不再用于存储服务。节点移除后，其上的数据会在其它节点重构。

在想要移除的节点一栏，单击“移除 Datanode”，完成了移除节点的操作。

3.1.3 删除域

操作：选择某个域，单击“删除”操作，可以完成域的删除工作。

注：当域有存储池的时候要先删除存储池才能删除域。

3.1.4 查询域

操作：查询框输入要查询的域名称(全称或者部分关键字)，支持模糊查询。



3.2 磁盘管理

3.2.1 卸载和挂载磁盘

- **卸载操作：**对此物理盘不纳入管理范围，此物理盘空间不作为系统的可用空间；

状态: GOOD-->OFFLINING-->OFFLINED(存储池界面磁盘状态同样的转变)

- **挂载操作：**将系统物理盘纳入管理范围，此物理盘的空间作为存储系统的空间，磁盘卸载后，其上的数据将发生迁移。

状态: OFFLINED<-->GOOD (存储池界面磁盘状态同样的转变)

3.3 存储池管理

ZettaStor DBS 软件通过在存储节点上部署轻量级的管理软件(DataNode)，把各个节点上的各类存储介质进行聚合，形成可统一管理的存储池，对外提供直接高性能的块设备接口。

存储池管理支持存储池的创建、删除、修改、存储池中磁盘信息查看、添加或者移除磁盘等操作。

3.3.1 新建存储池

操作：单击“新建”，进入新建存储池界面，如下图所示。

新建存储池

存储池名称: 存储池名称

类型: 容量

安全级别: 机柜级

描述: 描述

创建 取消

存储池名称：唯一标识，不能修改，不能重复；

类型：包括容量、性能、混合三个类型；

安全级别：包括机柜级和服务级；

描述：可输入对此存储池的简单描述，可写可不写；

单击“创建”，系统弹出存储池创建成功的信息。

3.3.2 修改存储池

操作：选中某个存储池，包括增加磁盘和移除磁盘，然后单击“修改”，系统跳到修改存储池界面，如图所示。

修改存储池

存储池ID: 3577288763904477229

存储池名称: storage-pool

类型: 容量

安全级别: 机柜级

描述: 描述

可添加的磁盘

| <input type="checkbox"/> | 实例ID | 主机IP地址 | 所在组编号 | 磁盘ID | 磁盘名称 | 逻辑空间 | 状态 |
|--------------------------|---------------------|------------|-------|---------------------|----------|-------|------|
| <input type="checkbox"/> | 2929984930422649389 | 10.0.1.159 | 3 | 513941098943397735 | /dev/sda | 9(GB) | GOOD |
| <input type="checkbox"/> | 1561961636391250889 | 10.0.1.157 | 2 | 1246347037949183415 | /dev/sda | 9(GB) | GOOD |
| <input type="checkbox"/> | 4412792139977965540 | 10.0.1.156 | 0 | 1150006413233770115 | /dev/sda | 9(GB) | GOOD |
| <input type="checkbox"/> | 4434995859500797341 | 10.0.1.158 | 1 | 1812643653415428398 | /dev/sda | 9(GB) | GOOD |

修改 取消

- 增加磁盘：形成可统一管理的存储池做为存储容量提供存储服务。

选中存储池名称中的磁盘，单击“修改”，系统弹出存储池修改成功的信息。

操作：单击存储池名称旁边的“+”，可看到如下信息：

| | | | | | |
|-------------------------------------|-------|----|------|----|----|
| <input type="checkbox"/> | 存储池名称 | 类型 | 安全级别 | 状态 | 描述 |
| <input checked="" type="checkbox"/> | vbn | 容量 | 机柜级 | 可用 | |

| 实例ID | 主机IP地址 | 所在组编号 | 磁盘ID | 磁盘名称 | 逻辑空间 | 状态 | 管理 |
|---------------------|------------|-------|---------------------|----------|-------|------|----------------------|
| 3756104078996979529 | 10.0.1.158 | 0 | 607086084681306055 | /dev/sda | 9(GB) | GOOD | 移除磁盘 |
| 69945187670361856 | 10.0.1.159 | 1 | 3405525218746628798 | /dev/sda | 9(GB) | GOOD | 移除磁盘 |
| 4533402302526432155 | 10.0.1.156 | 2 | 3198422940885315635 | /dev/sda | 9(GB) | GOOD | 移除磁盘 |
| 2068714593061261035 | 10.0.1.157 | 3 | 930227237184031126 | /dev/sda | 9(GB) | GOOD | 移除磁盘 |

- 移除磁盘：此磁盘不再用于存储服务。磁盘移除后，其上的数据会在其他节点重构。

操作：单击“移除磁盘”，就完成了移除磁盘的操作。

3.3.3 删除存储池

操作：选择某个存储池，单击“删除”操作，可以完成存储池的删除工作。

| <input type="checkbox"/> | 存储池名称 | 类型 | 安全级别 | 状态 | 描述 |
|--------------------------|-------|----|------|-----|----|
| <input type="checkbox"/> | vbn | 容量 | 机柜级 | 删除中 | |

3.3.4 查询存储池

操作：查询框输入要查询的存储池名称(全称或者部分关键字)，支持模糊查询。

查询存储池名称

3.4 卷管理

创建、修改域和存储池后，我们就能创建卷了。卷管理是客户应用的核心；提供对卷全面操作和状态呈现。功能包括创建卷，克隆卷、删除卷，扩展卷、修改卷的 QOS、挂载卷、卸载卷、卷访问规则设置、修改卷的缓存类型、转移卷及实时查看 segment Unit 状态列表。

3.4.1 管理卷界面

管理卷界面首页为卷列表，列举卷名称、卷大小、卷状态、所属域、所属存储池、精简配置、缓存类型、创建时间、卷详细信息。

| <input type="checkbox"/> | 卷名称 ▲ | 卷大小 | 卷状态 | 所属域 | 所属存储池 | 精简配置 | 缓存类型 | 创建时间 | 卷详细信息 |
|--------------------------|-------|---------|-----|--------|-------|------|------|---------------------|--|
| <input type="checkbox"/> | v1 | 10240.0 | 可用 | domain | pool | No | 内存缓存 | 2016/10/21 15:41:36 | 驱动 快照 详细 |

卷状态：分为正在创建、可用、不可用、正在删除、已删除。

所属域：可供此卷分配、迁移 segment 的 DataNode 范围；建议把具有相同读写速度的服务器划分在同一个域中；

所属存储池：可供此卷分配、迁移 segment 的存储磁盘范围；

精简配置：系统不是一次性把用户申请的存储空间分配到位，而是随着使用用户实际数据量的增长而逐步、少量地供给，从而有效提升存储资源利用效率。

缓存类型：分为内存缓存和 SSD 缓存；

卷详细信息：可查看 segment unit 状态及所在主机的关联列表，可对卷进行挂载驱动和卸载驱动，可对卷建立快照。

3.4.2 创建卷

操作：单击“卷-->新建卷”，进入创建卷界面，如图所示。

创建卷 ×

卷名

卷大小

精简配置

缓存类型

所在域

存储池

副本数量

卷名

卷大小

是

内存缓存

domain

pool

2

X1024M

创建 取消

卷名：卷的名称，唯一标识，不能修改，不能重复；

卷大小：为 segment 大小 1024M 的整数倍，segment 大小只在系统部署之

前允许修改；

精简配置：系统不是一次性把用户申请的存储空间分配到位，而是随着使用用户实际数据量的增长而逐步、少量地供给，从而有效提升存储资源利用效率。

缓存类型：分为内存缓存和 SSD 缓存；

所在域：卷存放的物理空间范围；

存储池：卷具体存在的物理磁盘；

副本数量：备份的数量，可以根据需要选择相应数量；

操作：单击“创建”后，系统跳转到到卷创建进度页面，如图所示。



当卷状态变为 Available 时，表示卷创建成功，可以进行挂载卷操作，选择相应的驱动类型，单击“挂载”，可以将卷挂载到某一个 DriverContainer 服务器上去。(在卷详细信息里也可完成此操作，效果一样)，如图所示。



挂载成功后，可以根据驱动的类型 PYD、ISCSI 的客户端工具连上卷。就可以像访问本地服务器硬盘一样对卷进行格式化、读写操作了。

3.4.3 配置卷访问规则

■ 卷访问规则管理

这部分的功能属于权限控制部分，包括权限定义及权限项和 segment 连接两部分。

操作：权限定义是全局的操作，单击“新建”，输入“远程主机 ip 地址”，和选择“读写权限”，然后单击“创建”，就完成了权限的定义。“读写权限”包括“Read-Only”和“Read/Write”，含义分别是“只读”和“读写”，如图所示。



■ 对卷设置访问规则

即应用已经创建好的客户端，只有获得授权的客户端才能访问存储资源。

操作：访问规则创建后，需要在卷页面，选中某个卷，单击“卷访问规则”，

选择要授权的客户端，单击“应用”，才能将权限项和卷联系起来，最终完成授权。



配置卷访问规则

卷名: v1

IP过滤器 提交

| <input type="checkbox"/> | 远端主机ip地址 | 读写权限 | 访问规则状态 |
|-------------------------------------|------------|------------|--------|
| <input checked="" type="checkbox"/> | 10.0.1.153 | Read/Write | 已授权 |
| <input type="checkbox"/> | 10.0.1.173 | Read/Write | 未授权 |

应用 取消

完成

注：每个卷最多应用两个卷访问规则。

3.4.4 配置卷缓存类型

操作：修改卷的缓存类型。



修改缓存类型

卷名 v1

缓存类型 内存缓存

修改 取消

可选择使用内存或者 SSD 来作为缓存，默认内存。

3.4.5 配置卷 QOS

■ 设置卷 QOS

限制卷某个时间段的 IOPS 和吞吐量，以保证关键应用的正常。



修改QOS

卷名 v1

➕新建

| 类型 | 开始时间 | 结束时间 | IOPS | 吞吐量 | 管理 |
|----|------|------|------|-----|----|
|----|------|------|------|-----|----|

完成

■ 新建动态 QOS

操作：选择类型“动态”，设置开始和结束时间，设置 IOPS 和吞吐量限制（大于 0 的数字），单击“添加”，如图所示。



The screenshot shows a dialog box titled "添加IO控制项" (Add IO Control Item). It contains the following fields:

- 类型 (Type):** A dropdown menu set to "动态" (Dynamic).
- 开始时间 (Start Time):** Two dropdown menus for hours and minutes, both set to "00".
- 截止时间 (End Time):** Two dropdown menus for hours and minutes, both set to "00".
- IOPS限制 (IOPS Limit):** Two input fields, both containing "无限制" (Unlimited).
- 吞吐量(KBPS) (Throughput (KBPS)):** Two input fields, both containing "无限制" (Unlimited).

At the bottom right, there are two buttons: "添加" (Add) in blue and "取消" (Cancel) in red.

■ 新建静态 QOS

操作：选择类型“静态”，设置开始和结束时间，设置 IOPS 和吞吐量限制（大于 0 的数字），单击“添加”，如图所示。静态类型的限制项会优先其他的限制项执行。



The screenshot shows a dialog box titled "添加IO控制项" (Add IO Control Item). It contains the following fields:

- 类型 (Type):** A dropdown menu set to "静态" (Static).
- IOPS限制 (IOPS Limit):** Two input fields, both containing "无限制" (Unlimited).
- 吞吐量(KBPS) (Throughput (KBPS)):** Two input fields, both containing "无限制" (Unlimited).

At the bottom right, there are two buttons: "添加" (Add) in blue and "取消" (Cancel) in red.

■ 修改 QOS

操作：在想修改的 QOS(包括动态和静态)一栏单击“修改”，类型、IOPS、吞



The screenshot shows a dialog box titled "修改IO控制项" (Modify IO Control Item). It contains the following fields:

- ID:** A text field containing "942293142".
- 类型 (Type):** A dropdown menu set to "静态" (Static).
- IOPS限制 (IOPS Limit):** Two input fields, the first contains "无限制" (Unlimited) and the second contains "1000".
- 吞吐量(KBPS) (Throughput (KBPS)):** Two input fields, the first contains "无限制" (Unlimited) and the second contains "5000".

At the bottom right, there are two buttons: "修改" (Modify) in blue and "取消" (Cancel) in red.

吐量分别修改为想要的值，单击“修改”。

■ 删除 QOS

操作：选择一个 QOS，单击“删除”，删除成功。

3.4.6 挂载驱动和连接、断开客户端

ZettaStor 系统支持 ISCSI、PYD 两种类型的驱动；

ISCSI：现有 SCSI 接口与以太网(Ethernet)技术结合，使服务器可与使用 IP 网络的储存装置互相交换资料。

PYD：让你可以将一个远程主机的磁盘空间，当作一个块设备来使用，就像一块硬盘一样使用它，你可以很方便的将另一台服务器的硬盘空间，增加到本地服务器上。

驱动状态包含 LAUNCHING、LAUNCHED 和 UNKNOWN(卷不可用时)。

■ PYD 单链路

操作：驱动类型选择 PYD，然后单击“挂载”，便可挂载驱动，驱动状态从“LAUNCHING”变为“LAUNCHED”卷挂载成功后如图所示。

驱动信息

卷名:v5

挂载 驱动类型 PYD 1 驱动用户数量 0

| 快照id | 驱动类型 | 驱动地址 | 驱动端口 | 驱动用户数量 | 驱动用户信息 | 驱动状态 | 驱动卸载 |
|-------|------|------------|------|--------|----------|----------|------|
| ----- | PYD | 10.0.1.156 | 1234 | 0 | clients: | LAUNCHED | 卸载 |

卸载所有驱动

完成

驱动挂载成功后，用户可以根据 DriverContainer 返回的挂载信息，连接客户端。

步骤如下：

- 1) 需要将 Pyd 所需要的文件放置到节点机器的/opt 目录下。具体来说，创建目录/opt/pyd/，将 nbd-client 和 pyd.ko 文件放入此目录下；
- 2) insmod pyd.ko (加载新内核)
- 3) 检查 pyd.ko 是否正确加载

```
# lsmod | grep pyd
```

```
# ls -l /dev/P*
```

```
# root @ vm53 in /opt/pytest/deploy/pengyun-deploy [16:5
$ lsmod | grep pyd
pyd                22094  0

# root @ vm53 in /opt/pytest/deploy/pengyun-deploy [16:5
$ ls -l /dev/P*
brw-rw----. 1 root disk 255,  0 Oct  8 14:45 /dev/Pyd0
brw-rw----. 1 root disk 255, 16 Oct  8 14:45 /dev/Pyd1
brw-rw----. 1 root disk 255, 32 Oct  8 14:45 /dev/Pyd2
brw-rw----. 1 root disk 255, 48 Oct  8 14:45 /dev/Pyd3
brw-rw----. 1 root disk 255, 64 Oct  8 14:45 /dev/Pyd4
brw-rw----. 1 root disk 255, 80 Oct  8 14:45 /dev/Pyd5
brw-rw----. 1 root disk 255, 96 Oct  8 14:45 /dev/Pyd6
brw-rw----. 1 root disk 255, 112 Oct  8 14:45 /dev/Pyd7
brw-rw----. 1 root disk 255, 128 Oct  8 14:45 /dev/Pyd8
brw-rw----. 1 root disk 255, 144 Oct  8 14:45 /dev/Pyd9
```

4) 将/dev/P*设备与主机连接(在连接客户端之前，首先对卷配置相应的访问规则)。命令如下：

```
# ./nbd-client 10.0.1.156 1234 /dev/Pyd*
```

```
# root @ vm53 in /opt/pyd [16:33:26]
$ ./nbd-client 10.0.1.156 1234 /dev/Pyd1
Negotiation: ..size = 10240MB
bs=1024, sz=10737418240 bytes
```

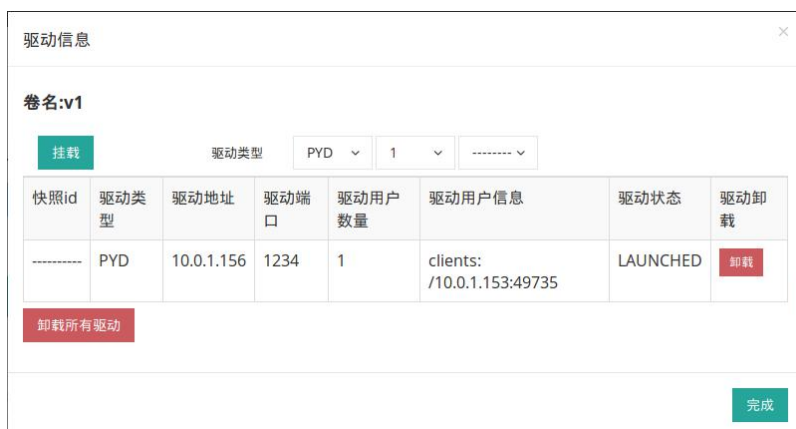
10.0.1.156 1234 为 PYD 服务端的 IP 和端口，即在 ZettaStor 的 Console 界面上挂载 PYD 时显示的 IP 地址和端口，bs 用于表示块大小，默认是 1024，可以是 512、1024、2048、4096，Pyd_device 映射到本地的哪个 Pyd 设备(如：/dev/Pyd1)。

5) 连接成功后，可以看到 nbd-client 进程。

```
# ps -ef | grep nbd
```

```
# root @ vm53 in /opt/pyd [16:36:00]
$ ps -ef | grep nbd
root    11648      1  0 16:34 ?        00:00:00 ./nbd-client 10.0.1.156 1234 /dev/Pyd1
root    11982 10207  0 16:36 pts/0    00:00:00 grep --color=auto --exclude-dir=.bzr --exclude-dir=CVS --exclude-dir=.git --e
xclude-dir=.hg --exclude-dir=.svn nbd
```

连接成功后，卷管理的页面显示如下：



接着可以挂载设备进行分区、格式化、创建文件系统和读写操作。

6) 分区：fdisk /dev/Pyd1 (以 Pyd1 为例)，然后使用命令“lsblk”查看后的结果如下：

```
Pyd1      255:16   0    6G   0 disk
└─Pyd1p1  255:17   0    6G   0 part
```

7) 格式化：mkfs.ext4 /dev/Pyd1p1(以 Pyd1 为例)

8) 创建一个挂载目录(随自己喜好)：mkdir /root/test-mount

9) 挂载：mount /dev/Pyd1p1 /root/test-mount

10) 卸载：umount /root/test-mount”

11) 使用命令 “./nbd-client -d /dev/Pyd*” (这里的 Pyd*与上面连接的地方要一致) 来断开客户端和服务端之间的连接。

```
# root @ vm53 in ~ [16:39:39] C:1
$ ./nbd-client -d /dev/Pyd1
Try opening nbd device
Done with opening nbd device
disconnect, sock, done
```

■ PYD 多链路

前提是部署系统的时候 DriverContainer 要部署 2 个以上。

操作：驱动类型选择 PYD，个数选择 2 或 3，然后单击“挂载”，便可挂载卷，驱动状态从“LAUNCHING”变为“LAUNCHED”卷挂载成功后如图所示。

驱动信息

卷名:v1

驱动类型 PYD 2

| 快照id | 驱动类型 | 驱动地址 | 驱动端口 | 驱动用户数量 | 驱动用户信息 | 驱动状态 | 驱动卸载 |
|-------|------|------------|------|--------|----------|----------|-----------------------------------|
| ----- | PYD | 10.0.1.158 | 1234 | 0 | clients: | LAUNCHED | <input type="button" value="卸载"/> |
| ----- | PYD | 10.0.1.156 | 1235 | 0 | clients: | LAUNCHED | <input type="button" value="卸载"/> |

1) 在 pyd 路径下配置 config/multipaths 文件，配置文件中 IP 地址和端口需要和 console 界面上显示的一致。

```
# This file contains all paths to send I/O request to server
# Each path is made up with host name and port.
# e.g.10.0.1.116 1234
10.0.1.156 1235
10.0.1.158 1234
```

2) 对卷配置相应的访问规则后用命令 “./nbd-client /dev/Pyd* -M” 来挂载，如图所示。

```
# root @ vm53 in /opt/pyd [16:S6:S7]
$ ./nbd-client /dev/Pyd1 -M
Negotiation: ..size = 10240MB
Negotiation: ..size = 10240MB
bs=1024, sz=10737418240 bytes
```

3) 挂载完毕后，在卷管理页面的显示如下：

驱动信息

卷名:v1

驱动类型 PYD 2

| 快照id | 驱动类型 | 驱动地址 | 驱动端口 | 驱动用户数量 | 驱动用户信息 | 驱动状态 | 驱动卸载 |
|-------|------|------------|------|--------|-------------------------------|----------|-----------------------------------|
| ----- | PYD | 10.0.1.158 | 1234 | 1 | clients: /10.0.1.153:37883 | LAUNCHED | <input type="button" value="卸载"/> |
| ----- | PYD | 10.0.1.156 | 1235 | 1 | clients: /10.0.1.153:57919 | LAUNCHED | <input type="button" value="卸载"/> |

过 30 秒左右变为下图所示的情况(热备方式)：

驱动信息

卷名:v1

挂载 驱动类型 PYD 2

| 快照id | 驱动类型 | 驱动地址 | 驱动端口 | 驱动用户数量 | 驱动用户信息 | 驱动状态 | 驱动卸载 |
|-------|------|------------|------|--------|----------------------------|----------|------|
| ----- | PYD | 10.0.1.158 | 1234 | 0 | clients: | LAUNCHED | 卸载 |
| ----- | PYD | 10.0.1.156 | 1235 | 1 | clients: /10.0.1.153:57893 | LAUNCHED | 卸载 |

卸载所有驱动

完成

此时，读写过程中其中一个链路中断，会自动切换到另外一个链路，读写不会中断。

可以使用命令“./nbd-client -d /dev/Pyd*”来断开客户端和服务端之间的连接。

■ ISCSI 单链路

操作：驱动类型选择 ISCSI，然后单击“挂载”，便可挂载卷，驱动状态从“LAUNCHING”变为“LAUNCHED”卷挂载成功后如图所示。

驱动信息

卷名:v5

挂载 驱动类型 ISCSI 1

| 快照id | 驱动类型 | 驱动地址 | 驱动端口 | 驱动用户数量 | 驱动用户信息 | 驱动状态 | 驱动卸载 |
|-------|-------|------------|------|--------|----------|----------|------|
| ----- | ISCSI | 10.0.1.156 | 3260 | 0 | clients: | LAUNCHED | 卸载 |

卸载所有驱动

完成

驱动挂载成功后，用户可以根据 DriverContainer 返回的挂载信息，对卷配置相应的访问规则后连接客户端。

在客户端机器上使用命令“iscsiadm -m discovery -p IP:Port -t sendtargets”来获取获取 iqn。然后使用命令“iscsiadm -m node -T iqn_number -p IP:Port -l”来挂载。


```
# root@vm53 in /opt/pyd [15:49:08]
$ iscsiadm -m discovery -p 10.0.1.156 -t sendtargets
10.0.1.156:3260,1 Zettastor.IQN:2358962684549075002
192.168.122.1:3260,1 Zettastor.IQN:2358962684549075002

# root@vm53 in /opt/pyd [15:57:44]
$ iscsiadm -m node -T Zettastor.IQN:2358962684549075002 -p 10.0.1.156 -l
Logging in to [iface: default, target: Zettastor.IQN:2358962684549075002, portal: 10.0.1.156,3260] (multiple)
Login to [iface: default, target: Zettastor.IQN:2358962684549075002, portal: 10.0.1.156,3260] successful.
```

挂载成功后，在 console 上可以看到：驱动用户数量变为 1，驱动用户信息显示正确。



通过命令 “ls -l /dev/disk/by-path/ip-*” 来查看挂载的设备

```
# root@vm53 in /opt/pyd [16:00:38]
$ ls -l /dev/disk/by-path/ip-*
lrwxrwxrwx. 1 root root 9 Oct 27 15:58 /dev/disk/by-path/ip-10.0.1.156:3260-iscsi-Zettastor.IQN:2358962684549075002-lun-0 -> ../../sdd
```

从这里可以知道，挂载的设备为/dev/sdd，这接着对设备进行分区（fdisk 命令）、格式化（mkfs 命令），挂载（mount 命令）之后就可以进行读写操作了。

读写完毕后先卸载（umount 命令），然后使用命令 “iscsiadm -m node -T iqn_number -p IP:Port --logout” 进行断开连接。

```
# root@vm53 in ~ [16:34:13]
$ iscsiadm -m node -T Zettastor.IQN:3676053741069835662 -p 10.0.1.156 --logout
Logging out of session [sid: 9, target: Zettastor.IQN:3676053741069835662, portal: 10.0.1.156,3260]
Logout of [sid: 9, target: Zettastor.IQN:3676053741069835662, portal: 10.0.1.156,3260] successful.
```

■ ISCSI 多链路

在使用 ISCSI 驱动多链路的时候，首先要在客户端安装 multipath-tools。

Ubuntu 环境下的安装命令如下：

```
sudo apt-get install multipath-tools
```

multipath 的配置文件为/etc/multipath.conf，multipath-tools 安装默认没有建立此文件，拷贝一个样本配置文件即可。使用命令：

```
cp /usr/share/doc/multipath-tools/examples/multipath.conf.synthetic
/etc/multipath.conf
```

默认的多路径设置是按照默认的参数配置的，这里需要根据我们的使用情况稍微调整一下，vi /etc/multipath.conf：

◆ 对于 defaults 设置：实际有用的就是下面这些参数：其余参数都可以注释掉：

```
defaults {
user_friendly_names yes
udev_dir /dev
path_grouping_policy failover
failback immediate
no_path_retry fail
}
```

```
defaults {
# udev_dir /dev
# find_multipaths yes
# polling_interval 10
# selector "round-robin 0"
# path_grouping_policy multibus
# path_grouping_policy failover
# getuid_callout "/lib/udev/scsi_id --whitelisted --device=/dev/%n"
# prio alua
# path_checker readsector0
# rr_min_io 100
# max_fds 8192
# rr_weight priorities
# failback immediate
# no_path_retry fail
# user_friendly_names yes
}
```

需要说明的有：

failback 设置 immediate 表示当前链路出现问题后立刻切换到下一条链路，这个需要与 open-iscsi 配置文件 /etc/iscsi/iscsid.conf 中的 node.session.timeo.replacement_timeout = 0 配合使用。

(重启服务 service iscsid restart，使得配置生效)

◆ multipath 配置在 datanode 节点时，需要注意 blacklist 配置，否则本地磁盘一旦被管理，会导致 datanode 识别不到磁盘，方法如下：

vi /etc/multipath.conf，添加如图内容：

```
blacklist {
wwid 26353900f02796769 #禁止管理本地的磁盘 (sd*)，建议使用wwid，防止盘符变化，导致错误
devnode "^(ram|raw|loop|fd|md|dm-|sr|scd|st)[0-9]*" #禁止管理ram,raw...等设备
devnode "^hd[a-z]" #禁止管理hd*光驱等设备
}
```

其中 wwid 通过以下命令获得：

```
scsi_id -g -u -s /block/sda
```

操作：驱动类型选择 ISCSI，个数选择 2 或 3，然后单击“挂载”，便可挂载卷，驱动状态从“LAUNCHING”变为“LAUNCHED”卷挂载成功后如图所示。

驱动信息

卷名 v5

挂载

驱动类型

ISCSI

2

| 快照id | 驱动类型 | 驱动地址 | 驱动端口 | 驱动用户数量 | 驱动用户信息 | 驱动状态 | 驱动卸载 |
|-------|-------|------------|------|--------|----------|----------|------|
| ----- | ISCSI | 10.0.1.156 | 3260 | 0 | clients: | LAUNCHED | 卸载 |
| ----- | ISCSI | 10.0.1.158 | 3260 | 0 | clients: | LAUNCHED | 卸载 |

卸载所有驱动

完成

对卷配置相应的访问规则后，在客户端机器上使用命令“iscsiadm -m discovery -p IP:Port -t sendtargets”来获取获取 iqn。

```
# root @ vm53 in /opt/pytest/deploy/pengyun-deploy [16:23:21]
$ iscsiadm -m discovery -p 10.0.1.156 -t sendtargets
10.0.1.156:3260,1 Zettastor.IQN:2358962684549075002
192.168.122.1:3260,1 Zettastor.IQN:2358962684549075002

# root @ vm53 in /opt/pytest/deploy/pengyun-deploy [16:38:26]
$ iscsiadm -m discovery -p 10.0.1.158 -t sendtargets
10.0.1.158:3260,1 Zettastor.IQN:2358962684549075002
192.168.122.1:3260,1 Zettastor.IQN:2358962684549075002
```

(从这里可以看出它们的 iqn 都是一样的)

然后使用命令“iscsiadm -m node -T iqn_number -p IP:Port -l”来连接服务器和客户端。

这时 console 上的显示如下图所示(与 NBD 多链路有所不同)

驱动信息

卷名:v5

挂载

驱动类型

ISCSI

2

| 快照id | 驱动类型 | 驱动地址 | 驱动端口 | 驱动用户数量 | 驱动用户信息 | 驱动状态 | 驱动卸载 |
|-------|-------|------------|------|--------|---------------------|----------|------|
| ----- | ISCSI | 10.0.1.156 | 3260 | 1 | clients: 10.0.1.153 | LAUNCHED | 卸载 |
| ----- | ISCSI | 10.0.1.158 | 3260 | 1 | clients: 10.0.1.153 | LAUNCHED | 卸载 |

卸载所有驱动

完成

使用命令 iscsiadm -m node 查看映射结果

```
# root @ vm53 in /opt/pytest/deploy/pengyun-deploy [16:
$ iscsiadm -m node
10.0.1.156:3260,1 Zettastor.IQN:2358962684549075002
10.0.1.158:3260,1 Zettastor.IQN:2358962684549075002
192.168.122.1:3260,1 Zettastor.IQN:2358962684549075002
```

通过命令 “multipath -ll” 查看多路径拓扑信息

```
# root @ vm53 in /opt/pytest/deploy/pengyun-deploy [16:49:48]
$ multipath -ll
Oct 27 16:50:53 | multipath.conf line 31, duplicate keyword: defaults
Oct 27 16:50:53 | multipath.conf +32, invalid keyword: udev_dir
mpathd (1494554000000000f9e7cbecf842b860c42ac81777092c9c) dm-2 IET ,VIRTUAL-DISK
size=12G features='0' hwhandler='0' wp=rw
|+- policy='service-time 0' prio=1 status=active
| - 5:0:0:0 sdd 8:48 active ready running
|+- policy='service-time 0' prio=1 status=enabled
| - 6:0:0:0 sde 8:64 active ready running
```

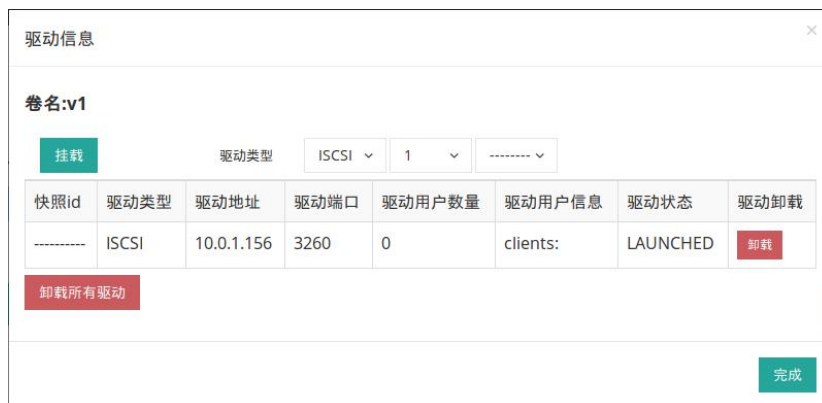
(如果结果显示中 2 个链路的状态都为 active，这说明 2 条链路都可以传输数据，需要检查一下多链路配置文件中 path_grouping_policy 的配置)

通过命令 “ls -l /dev/mapper” 查看 map 设备。

```
# root @ vm53 in /opt/pytest/deploy/pengyun-deploy [16:50:53]
$ ls -l /dev/mapper
total 0
lrwxrwxrwx. 1 root root 7 Oct 8 14:45 centos-root -> ../dm-1
lrwxrwxrwx. 1 root root 7 Oct 8 14:45 centos-swap -> ../dm-0
crw----- 1 root root 10, 236 Oct 8 14:45 control
lrwxrwxrwx. 1 root root 7 Oct 27 16:40 mpathd -> ../dm-2
```

操作完毕之后，使用命令 “iscsiadm -m node -T iqn_number -p IP:Port --logout” 进行断开连接。

3.4.7 卸载驱动



操作：驱动挂载成功后，可看到“卸载”按钮，单击“卸载”按钮，便可进行驱动的卸载，挂载信息没有了。

3.4.8 查询卷

Admin 用户可以查询所有的卷，Regular 只能查询自己创建的卷。

操作：单击“卷”，进入管理卷页面，页面右上角有查询卷功能。打开页面时系统自动列出该类型用户所能查询到的所有卷。

查询卷名称

3.4.9 卷详细信息

实时查看卷包含的所有 segment 的数量和状态。Segment 的运行状态包括：OK、Unknown、Transit、deleting。

当卷进入稳定状态时，所有的 segment 是 OK 状态。如果服务器的物理磁盘发生增减，segment 将进行系统的再平衡，迁移到新的空间。此时有“Unknown”和“Transit”两种状态。Unknown 表示目前系统未得到当前 segment 状态响应，系统准备将此 segment 迁移到其他物理条件良好的空间。Transit 表示 segment 正在迁移。通过状态查询的下拉框里选择 ALL 可以查询到当前所有 segment，选择 STABLE 可以查询到当前所有 STABLE 状态的 segment，选择 AVAILABLE 可以查询当前所有可用状态的 segment，选择 UNAVAILABLE 可以查询当前所有不可用状态的 segment。

3.4.10 删除卷

操作：选择某个卷，单击“删除”按钮，可以进行卷的删除操作。

3.4.11 回收卷

执行卷的删除操作后，卷的状态很快就变成 Deleting，并且持续一段时间才变成 Deleted。segment 变成 deleting 状态的时候，是可以进行数据回收操作的。但是在 segment 从 deleting 状态变成 deleted 状态的最后 1 分钟，是不能进行数据回收的。

| | | | | | | | |
|--------------------------|----|--------|---------|--------|------|----|---|
| <input type="checkbox"/> | v1 | 5120.0 | Deleted | domain | pool | No | 驱动 快照 回收 详细 |
|--------------------------|----|--------|---------|--------|------|----|---|

操作：单击“回收”，就完成了卷的回收操作。

3.4.12 扩展卷

创建卷完成后，如果需要增加卷的大小，可以通过扩展卷界面进行扩展。

操作：如图所示，填写“扩展大小”，单击“扩展”就完成了卷的扩展操作。扩展成功后可通过“管理卷”界面中的“卷大小”可以查看到卷扩展的结果。（即卷大小增加了刚刚的扩展大小）



3.4.13 迁移卷

操作：我们有时需要将卷迁移到其他域的存储池；

迁移前后卷的信息不会有改变。



3.4.14 卷的克隆

克隆是指基于某个卷，创建出一个新的卷。新的卷不仅在容量上与作为基线的卷保持一致，而且其上的数据也是由作为基线的卷拷贝而来。另外，克隆还可以基于卷的某个快照来进行，方便对 volume 的某历史时刻的数据进行拷贝。

克隆出来的卷和作为基线的卷没有区别，任何普通的卷操作皆可应用于被克隆出来的卷。

操作：单击“克隆”按钮，选择克隆源，可以对该卷进行克隆，如图所示。

3.4.15 卷的快照

卷创建后用户随时可以创建快照，保存卷在某一个时刻的状态和数据。后续用户想要恢复到历史某个时刻的数据，可以将卷回滚到在那个时间创建的快照。

■ 创建快照

操作：依次输入快照名字(必须写)和详细描述(可写可不写)，单击“创建”，创建快照成功，如图所示。

| 快照ID | 快照名字 | 详细描述 | 创建时间 | 快照状态 | 管理 |
|------|------|------|------------------------|-----------|---------------------------------|
| 1 | s1 | | 10/25/2016, 3:20:31 PM | Available | 删除 回滚 |

■ 回滚快照

将卷回滚到指定的快照，回滚结束后，卷的数据将恢复到快照创建时的数

| 快照ID | 快照名字 | 详细描述 | 创建时间 | 快照状态 | 管理 |
|------|------|------|------------------------|-----------|---------------------------------|
| 1 | s1 | | 10/25/2016, 3:20:31 PM | Available | 删除 回滚 |

快照回滚中...

据；

操作：单击“回滚”，提示“快照回滚中...”。

系统支持单卷上创建的 snapshot 数量为 128 个。快照创建得越多，占用的磁盘空间越多。快照创建后对某块数据进行写操作时，将发生 COW(Copy On Write)，即新分配一块空间保留快照创建时的数据。当快照回滚操作完成后，回滚对应的快照，而其以后的快照都会被删除。

✧ 实例

在 volume1 上创建了 snapshot1～snapshot10，在 snapshot5 上回滚完成后，snapshot5～snapshot10 将被删除，只有 snapshot1～snapshot4 被保留。

■ 删除快照

操作：单击“删除”，此快照就从页面删除。

✧ 实例

在 volume1 上创建了 snapshot1～snapshot10，删除 snapshot5 后，snapshot5～有 snapshot1～snapshot4、snapshot6～snapshot10 还在列表里。(这点和回滚快照不一样)

3.4.16 孤儿卷

孤儿卷的主要作用是提醒用户此卷不可恢复使用，可删除。

卷由若干个 segment 组成，每个 segment 正常情况下由 3 个 segment unit 组成，当某个卷的所有 segment 都只有一个 segment unit 时，此卷就是孤儿卷。

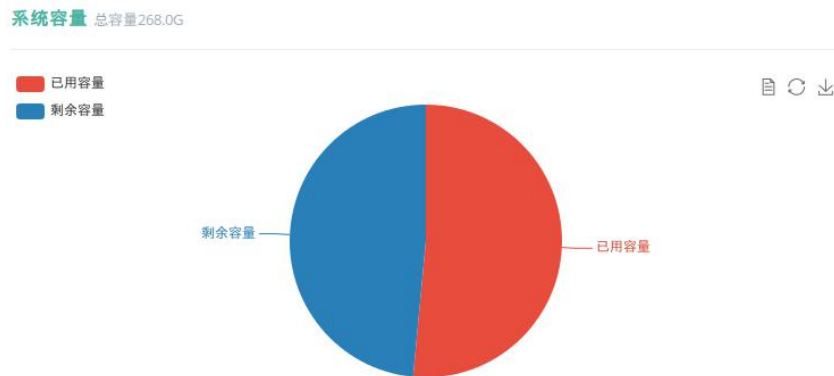
第四章 系统管理与监控

4.1 主页

实时动态的展现系统级容量参数、性能参数和状态参数。刷新时间间隔为 5s/10s/30s/60s。系统页面显示：包括系统容量、IOPS、吞吐量、客户端、用户、卷、节点以及系统容量趋势显示。

- 通过饼图直观展现系统容量的使用大小及百分比，红色为已用容量，蓝色

为剩余容量。



- 通过柱状图展现系统容量的使用情况，藏青色为总容量，红色为已用容量，蓝色为剩余容量。



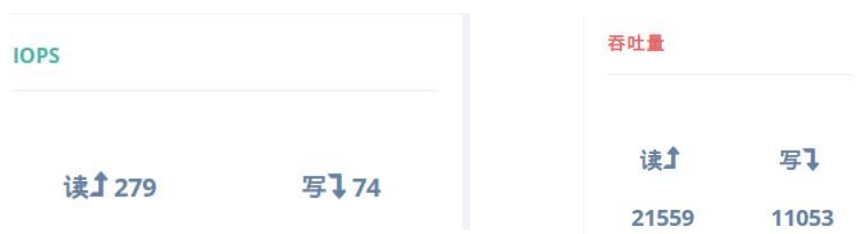
- 当前系统已经创建的卷的数量以及各状态卷的数量。 为正常状态的卷的数量； 为警告状态的卷的数量； 为错误状态卷的数量。



- 当前系统的 DataNode 节点总数以及各状态节点的数量。✅ 为正常状态的节点的数量；⚠️ 为警告状态的节点的数量；❌ 为错误状态节点的数量。



- 系统所有卷当前读写的性能数据之和，包括 IOPS 和吞吐量两个参数。



- 已连接卷的客户端总数为 0。



- 当前登陆的用户：admin。

4.2 系统服务管理

4.2.1 服务管理界面

向用户展示系统当前所有服务的名称、所在主机、端口状态及服务状态。包 DIH、InfoCenter、ControlCenter、MonitorCenter、DriverContainer、DataNode 六项服务。在该页面根据输入的组 ID 和服务名定位服务，如果不输入，则查询所有的服务。

| <input type="checkbox"/> | 服务名称 | 组ID | 所在主机 | 端口 | 服务状态 |
|--------------------------|-----------------|-----|------------|-------|------|
| <input type="checkbox"/> | ControlCenter | | 10.0.1.155 | 8010 | 正常 |
| <input type="checkbox"/> | DataNode | 2 | 10.0.1.156 | 10011 | 正常 |
| <input type="checkbox"/> | DataNode | 0 | 10.0.1.158 | 10011 | 正常 |
| <input type="checkbox"/> | DataNode | 3 | 10.0.1.157 | 10011 | 正常 |
| <input type="checkbox"/> | DataNode | 1 | 10.0.1.159 | 10011 | 正常 |
| <input type="checkbox"/> | DIH | | 10.0.1.158 | 10000 | 正常 |
| <input type="checkbox"/> | DIH | | 10.0.1.156 | 10000 | 正常 |
| <input type="checkbox"/> | DIH | | 10.0.1.157 | 10000 | 正常 |
| <input type="checkbox"/> | DIH | | 10.0.1.159 | 10000 | 正常 |
| <input type="checkbox"/> | DIH | | 10.0.1.155 | 10000 | 正常 |
| <input type="checkbox"/> | DIH | | 10.0.1.154 | 10000 | 正常 |
| <input type="checkbox"/> | DriverContainer | | 10.0.1.155 | 9000 | 正常 |
| <input type="checkbox"/> | DriverContainer | | 10.0.1.156 | 9000 | 正常 |
| <input type="checkbox"/> | InfoCenter | | 10.0.1.155 | 8020 | 正常 |

服务状态有 6 种：正常、挂起、停止、失败、丢失、未知。

4.2.2 服务管理操作

系统提供的服务功能：可以启动、停止、移除某服务。

停止：首先选中某个服务，单击“停止”，状态变化：正常或挂起-->停止；

启动：首先选中某个服务，单击“启动”，状态变化：停止-->正常或挂起；

移除：首先选中某个服务，单击“移除”，状态变化：正常或挂起-->失败-->丢失，然后从界面上消失。

由于 DIH 服务的功能是获得和传递各项服务的状态；所有当停止 DIH 服务后，其所在服务器的各项服务器变成“未知”的状态。

| | | | | | |
|--------------------------|---------------|--|------------|------|----|
| <input type="checkbox"/> | ControlCenter | | 10.0.1.157 | 8010 | 丢失 |
|--------------------------|---------------|--|------------|------|----|

4.3 系统性能监控与管理

4.3.1 创建性能任务

创建性能任务是从各个角度来观察我们的系统的日常运转。性能任务的状态包括 3 种：新建、停止、运行中。

操作：单击“新建”，就能到达新建性能任务界面，如图所示。

任务名称：任务的名称，唯一标识，不能修改，不能重复；(必须输入)

任务描述：对该任务的简单描述；(可写可不写)

单击“继续”，到达选择性能项界面，性能项分为三类：MACHINE、VOLUME、STORAGE_POOL；

- ❖ **MACHINE：**包括所有的服务节点，其中每个节点包含 12 个性能项：网络下行错误计数、实体机使用量、网络下行包速率、实体机 CPU 空闲百分比、网络下行丢包、网络上行包速率、网络上行错误计数、网络上行丢包、实体机内存总量、网络上行数据速率、网络下行数据速率、实体机剩余内存。
- ❖ **VOLUME：**包括所有卷，其中每个卷包含 4 个性能项：对卷的写总大小、对于卷的写 IO 计数、对卷的读总大小、对于卷的读 IO 大小。
- ❖ **STORAGE_POOL：**包含所有存储池，其中每个存储池包含 1 个性能项：存储利用率。

单击“继续”，到达设定时间范围界面；可以设置想要的时间区间，若想多个时间区间都执行该性能任务，可以单击“添加时间区间”；频率：必填，单位为 s；

单击“继续”，到达确认信息页面；该页面显示之前三个页面所选择和输入的信息：创建新任务、选择性能项和设定时间范围的信息。

单击“提交”，性能任务创建完成，以下是性能列表：显示的信息与创建时所输入信息一致；

| <input type="checkbox"/> | 任务名称 | 任务描述 | 任务频率 | 任务状态 | 任务执行时间 |
|--------------------------|--------------------------|--|------|------|-------------------|
| <input type="checkbox"/> | Default performance task | Pull CPU,memory,network performance data | 5 | 新建 | 15:40:11~17:03:31 |
| <input type="checkbox"/> | fio | | 1 | 运行中 | 00:00:00~23:55:00 |

4.3.2 启动性能任务

操作：选中要启动的性能任务，单击“启动”，这个性能任务就会生效。

4.3.3 停止性能任务

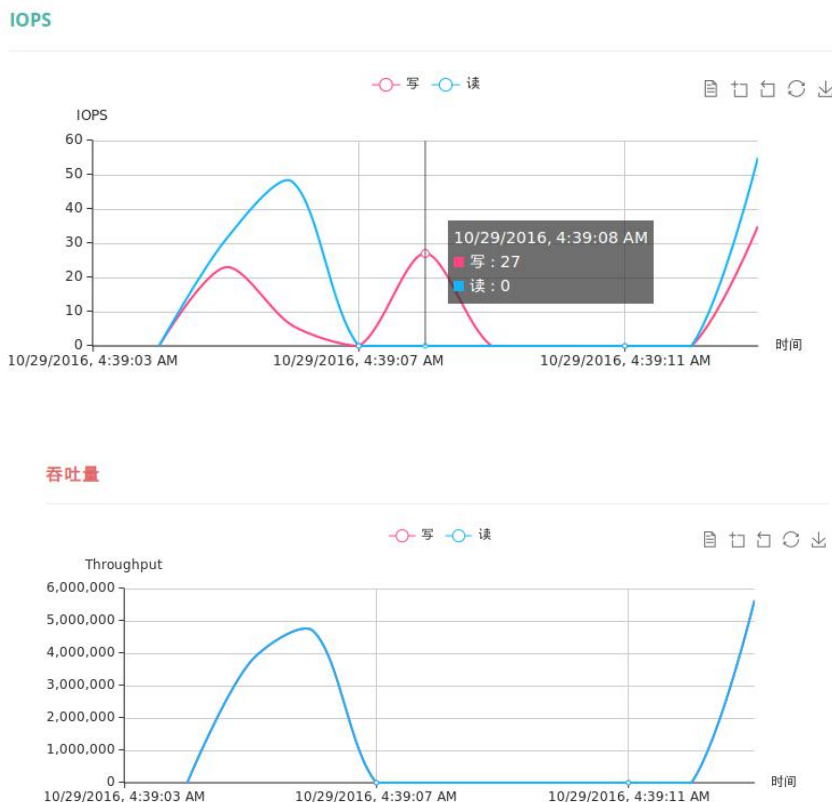
操作：选中要停止的性能任务，单击“停止”，这个性能任务就会失效。

4.3.4 删除性能任务

操作：选中要删除的性能任务，点单击“删除”，这个性能任务就会从界面上删除。

4.3.5 利用性能任务监控卷性能

创建性能任务：性能项选择 VOLUME: v1 的 4 个性能指标，并启动这个性能任务，并对卷进行 fio 读写，此时在主页-卷界面，选择创建性能任务时选择的卷，界面会实时展现各卷的读写性能，包括如下四个指标：卷的 IOPS、吞吐量、健康指数和延迟。





健康指数：此指标显示该卷的健康状态。显示绿色表示该卷健康状态良好；显示红色表示该卷不可用；黄色表示该卷可用，但部分 segment unit 状态异常。

4.3.6 利用性能任务监控实例性能

创建性能任务：性能项选择 MACHINE: 某个服务节点的 12 个性能指标，并启动这个性能任务。在主页-实例信息页面，选择刚刚创建性能任务时选择的节点，即实例，可看到该节点的 CPU 使用率和内存使用率。



- 对于其他的性能项可用同样的方法来新建，并启动、停止、删除性能任务。

4.4 数据重构与再平衡

4.4.1 数据重构

当存储单元(节点或者磁盘)发生故障或者被删除后，被删除存储单元上的数据会在其他健康存储单元上进行自动重构(rebuild)，选择重构的存储单元本着负载轻优先的原则，确保数据的负载均衡。Rebuild 期间不会影响存储业务的正常访问。重构的存储空间为存储系统空闲的存储空间，不必单独分配 rebuild 存储空间。

4.4.2 数据负载均衡

当一个新的存储单元(磁盘或者节点)加入后，系统会自动识别和接纳，并且自动将系统中的数据向新的存储单元进行平衡(rebalance)，确保数据负载恢复到一个平衡状态。平衡完成后，每个卷的数据都会均匀分布在所有(包括新加入)存储单元上，这样卷的性能得到最大化的发挥。Rebalance 期间不会影响存储业务的正常访问。Rebalance 可以人工触发也可以自动触发。可以根据 rebalance 的速度和和对业务的影响程度定制不同级别的 rebalance 规则供用户选择。

4.5 系统告警监控与管理

该功能是在系统出现异常情况的时候，比如网络延迟、网络断开、驱动状态不好等异常信息呈现出来，以便及时发现和处理，更好的维护系统。

4.5.1 查询告警

操作：输入告警名称时只显示该告警，不输入时显示全部告警；

告警有 4 个等级：Critical、Major、Minor、Warning；All 表示显示全部告警。

| | | | | |
|---|------------|--------------|-------|---|
| All | 移除 | 模板 | 3s | 刷新 |
| All | 查询告警名称 | | | |
| Critical | 对象 | 告警名称 | 告警级别 | 描述 |
| Major | 1.155:3260 | driver-alarm | MAJOR | Driver on volume 10.0.1.155:3260 is in bad status |
| Minor | | | | |
| Warning | | | | |
| 显示结果, 共 1 项 | | | | |
| <div> <div>«</div> <div><</div> <div>1</div> <div>></div> <div>»</div> </div> | | | | |

4.5.2 移除告警

操作：选中要移除的告警，单击“移除”，就能移除该告警。

4.5.3 告警模板管理

我们可以创建告警模板来设置什么情况下产生告警，当告警条件达到我们设置的标准就会产生告警信息，从不同角度来监控我们的系统，更好的维护我们的系统。

■ 新建告警模板

操作：单击“新建”，进入新建告警页面，如图所示。

创建告警模板

名称 *

名称

等级 *

Critical

描述

描述

建议

告警条件 *

时间窗

单位: s

持续时间

单位: s

系统卷数量

>

次数

次数

+添加告警条件

创建

取消

名称：告警的名称，唯一标识，不能修改，不能重复；(必须输入)

等级：包括 Critical、Major、Minor、Warning 四个；

描述：对该告警的简单描述，可写可不写；

建议：可写可不写。

◆ **告警条件：**包括时间窗、持续时间和告警条件发生的次数。

(注：可添加多个告警条件。)

- 1) 时间窗：我们定义一个时间长度为时间窗；
 - 2) 持续时间：时间窗滑动从开始时刻到结束时刻之间的时间长度为持续时间；
 - 3) 告警条件包括以下 23 种：系统卷数量、网络下行错误计数、实体机使用量、网络下行包速率、对卷的写总大小、实体机 CPU 空闲百分比、存储池利用率、对于卷的写 IO 计数、网络下行丢包、网络上行包速率、网络上行错误计数、JVM 最大内存量、JVM CPU 使用量、JVM 已提交内存、网络上行丢包、实体机内存总量、网络上行数据速率、网络下行数据速率、对卷的读总大小、JVM 预设内存上限、实体机剩余内存、JVM 已使用内存量、对于卷的读 IO 计数。
- 设置告警条件发生的次数后单击“创建”，告警模板创建成功。

■ 移除告警模板

操作：当你不需要这个告警模板时，选择某个告警模板，单击“移除”，该告警模板就从界面上删除了；

4.5.4 通过邮件发送告警

■ SMTP 服务器管理

遵循 SMTP 协议的发送邮件服务器，用来发送或中转发出的电子邮件。

操作：

- 添加、修改 SMTP 服务器管理；
- 1) 添加 SMTP 服务器管理界面如图所示。

添加SMTP服务器

名称

IP

端口

后缀

验证 True ▼

SSL True ▼

添加 取消

2) 修改 SMTP 服务器管理界面如图所示。

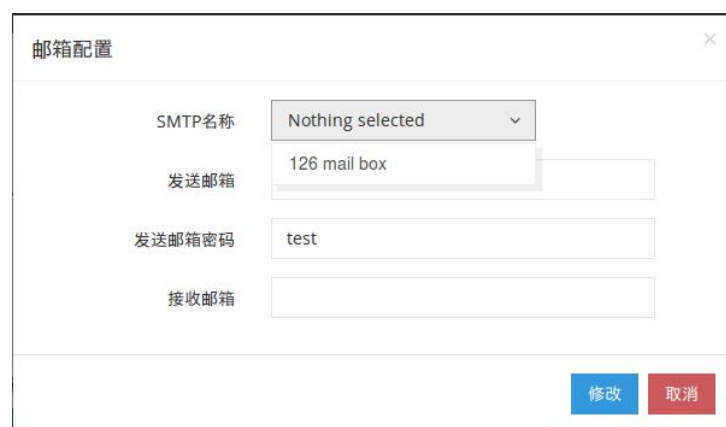


- 删除 SMTP 服务器管理：直接单击“删除”，此 SMTP 服务器管理就从界面上删除成功。

■ 邮箱配置

产生告警后会把这个信息以邮件的形式发送给管理员等。这里是对某个 SMTP 服务器管理进行发送邮箱和接受邮箱的设置。

操作：单击“修改”，邮箱配置成功。



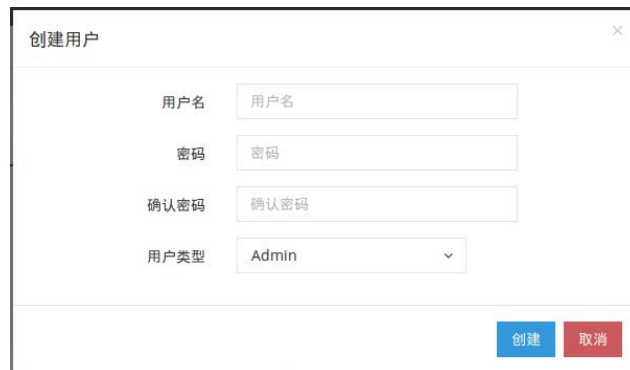
4.6 用户管理

系统登录 console 需通过用户名和密码验证、鉴权。系统提供三类账户管理功能，分别为 Superadmin、Admin 和 Regular 三种类型。Superadmin 是系统保

留账户 **admin** 的权限类型，能够看到所有的用户信息；**Admin** 是管理员用户，系统最高权限，可以查看、管理所有的用户创建的卷。**Regular** 是一般权限用户，只能查看、管理自己创建的卷。

操作：

- 创建用户：依次输入用户名，密码，确认密码，并选择用户类型，便可完成用户的创建。



创建用户对话框，包含以下输入项：

- 用户名：输入框，显示“用户名”
- 密码：输入框，显示“密码”
- 确认密码：输入框，显示“确认密码”
- 用户类型：下拉菜单，显示“Admin”

底部有两个按钮：创建（蓝色）和取消（红色）。

- 删除用户：当需要删除账号时，可以选择要删除的用户，单击“删除”按钮。只有 **Admin** 用户才能删除其他账号。
- 重置密码：当忘记账号密码时，可请求 **Admin** 用户重置密码。单击“重置密码”，进入密码重置页面。系统列出了系统已有的所有用户，用户比较多时，也可以在该页面根据用户名查询到对应的用户。选择要重置的用户所在的行，单击“重置密码”，便可完成密码的重置，重置后的密码为 **pengyun**。
- 修改密码：单击“修改密码”，进入密码修改页面，如图所示。



修改密码对话框，包含以下输入项：

- 旧密码：输入框，显示“旧密码”
- 新密码：输入框，显示“新密码”
- 确认新密码：输入框，显示“确认新密码”

底部有两个按钮：修改（蓝色）和取消（红色）。

一系列用户操作后的用户列表如图所示。

+新建
删除
重置密码
修改密码
3s
刷新

5
查询用户名

| <input type="checkbox"/> | 用户名 | 角色 |
|--------------------------|-------|------------|
| <input type="checkbox"/> | admin | SuperAdmin |
| <input type="checkbox"/> | pool | Regular |

显示第 1 至 2 项结果, 共 2 项
<<
<
1
>
>>

4.7 许可证管理

license 管理(试用版本无需做此操作)

- 1) 单击“生成序列号”，将系统生成的序列号提交给项目实施现场责任人，现场责任人向鹏云专职人员申请许可证后提交给客户。
- 2) 客户将许可证代码粘贴到许可证第三个输入框，单击“更新许可证”，许可证生效。

详细信息

生成序列号

更新许可证

| 类型 | 试用期(天) | 剩余试用期(天) | 功能令牌 | | 资源令牌 | |
|-------|--------|----------|------|----|------|--------------|
| Trail | 30 | 30 | 类型 | 值 | 类型 | 值 |
| | | | ALL | ON | | CAPACITY1024 |